DEPARTMENT OF HEALTH & HUMAN SERVICES PUBLIC HEALTH SERVICE

Memorandum

Food and Drug Administration
Center for Biologics Evaluation and Research
1401 Rockville Pike
Rockville, MD 20852

Division of Clinical Trial Design and Analysis
HFM-576

Date: October 1, 1999

From: Marc Walton, MD, PhD; OTRR/DCTDA

Subject: Medical Officer Review

Through: Karen Goldenthal MD, Director, OVRR/DVRPA

BLA 98-1396

Elan Pharmaceuticals, Inc.

Botulinum Toxin Type B

for Treatment of Cervical Dystonia

Clinical Review

OVERVIEW

Athena Neurosciences, Inc. submitted a marketing application, BLA 98-1396 on December 22, 1998 for botulinum toxin type B for use in the treatment of cervical dystonia. Athena was subsequently acquired by Elan Pharmaceuticals, Inc. a all rights and ownership for the product transferred to Elan. The entire clinical development program of the product was conducted by Athena. In order to avoid confusion for future readers of this document, this review will largely refer to Elan as the developer, manufacturer, and marketer of the product, irrespective of whether or not study conduct or other events occurred prior to acquisition of Athena by Elan Pharmaceuticals.

Elan has proposed the trade name NeuroBloc for the product. The proposed use of the product is in the treatment of cervical dystonia, with the proposed indication for labeling stated as: "NeuroBloc is indicated for the treatment of patients with cervical dystonia. In addition, NeuroBloc is indicated for the management of cervical dystonia in patients who ——— ——— . NeuroBloc treatment reduces the pain, ——— ', and severity of dystonia, ———

Scope of the Review

An extensive clinical development program was conducted. There are 11 studies presented in the application. Three were phase 1 open label studies, two phase 1 pharmacology studies in normal volunteers. There were six studies with sizable numbers of subjects in phase 2 and 3 studies. All studies were conducted under US IND, BB-IND ———

With the exception of important safety issues, this review document is limited to review of the six clinical studies that have occurred in the later stages of clinical development. This includes two dose ranging phase 2 randomized and controlled studies, two phase 3 randomized, controlled confirmatory studies, and two open label safety studies. The earlier studies have been incorporated only in the inclusion of adverse events in Integrated Summary of Safety component of review and a brief synopsis of these provided in an Appendix.

At the time of submission of the application, the two open label safety studies were still ongoing, and only an interim summary report was supplied from these studies.

Late in the review period, September 1999, Elan submitted a safety update that included additional information from the ongoing open label safety studies. Because of the lateness of this information in the review cycle, this submission has not been incorporated into the review.

TABLE OF CONTENTS

INTRODUCTION

Botulinum Toxins

Clostridia botulinum is an anaerobic bacterium that produces a neurotoxin. The clinical disease botulism and the association with tainted foods was described in the late 18th-early 19th centuries, from whence the name derives; German blood sausage was the most highlighted source, and *botulus* the Latin for sausage. The organism was first isolated and identified by E. van Ermengem in the late 19th century and shown to produce a toxin. Different strains of C. botulinum have been identified to produce 7 different types of this neurotoxin, designated types A through G, initially distinguished largely as different serotypes. C. butyricum produces a neurotoxin similar to type E, and C. baratii a toxin similar to type F (but neither are identical). Botulinum toxin has been called the "most poisonous poison", and thought to be the most lethal substance known (on a per molecule or per weight basis). Types A, B, and E are thought to account for the great majority of cases of human poisoning. The toxin is heat labile.

The toxin proteins of all types are synthesized by the bacillus as a single chain polypeptide of MW approximately 150 kD. A selective proteolysis step (nicking) is important in achieving the fully active toxin molecule with two chains (designated the heavy and the light chain) joined by disulfide bonds.

The toxins are all zinc metalloproteases, and their mechanism of toxicity is generally similar. All bind at specific receptor sites on cholinergic presynaptic terminals, although the specific receptor sites do not appear to be identical for all of the serotypes. The toxins are taken up by endocytosis, and in a process dependent upon a transmembrane pH difference, form pores in the endocytic vesicle membrane through which the light chain (possibly with some of the heavy chain) translocates into the cytosol. Once in the cytosol the protease toxin is active and each cleaves a specific protein critically involved in the neurotransmitter vesicle release process. There are three synaptic terminal proteins identified which have the specific proteolytic site for toxin types A through F (toxins A, E on SNAP-25, B,D,F,G on Synaptobrevin [VAMP], and type C on Syntaxin). Thus, toxin effects result in the failure of transmission at the neuromuscular junction. The toxin's effects are most prevalent at the neuromuscular junction, but have some effects also at the autonomic cholinergic terminals. Central nervous system botulinum toxin effects are not prominent when toxin is administered in the periphery, but in vitro the toxin is active in synaptosome preparations from CNS.

Consistent with the loss of neuromuscular junction transmission, clinical botulism consists of a flaccid paralytic disease. Symptoms stemming from ocular muscle paralysis are often the first noticed. The most threatening aspects are the loss of pharyngeal and diaphragmatic muscle function leading to risks of respiratory failure and aspiration, and are the primary causes of death in botulism patients. There is no known mechanism to reverse the synaptic failure once it has occurred. Reversal of the paralysis occurs only through natural means, and appears to occur by sprouting of new nerve terminals.

Botulinum toxin functions by producing a state of functional denervation. While immediately the nerve endings remain juxtaposed to the muscle (the neuromuscular junction, NMJ), neuromuscular transmission failure occurs, and the muscle responds as if actually denervated. Acetylcholine receptors begin to appear in a diffuse pattern along the entire muscle fiber, not localized to the NMJ. The nerve however, remains viable, and responds with sprouting of new nerve terminals. These form new neuromuscular junctions with the muscle fibers. Several months can be required to develop adequate sprouting to replace the disabled terminals. This sprouting and formation of new NMJ is responsible for the recovery of muscle function, and the consequent loss of benefit of the toxin over the subsequent months.

The type A neurotoxin is presently the only type in widespread clinical use and commercially available. However, several other serotypes are under investigation by various manufacturers. The Type B toxin developed by Elan Pharmaceuticals (beginning as Athena Neurosciences prior to acquisition of Athena by Elan) is the focus of this marketing application.

Potency of botulinum toxins is generally assessed with a mouse intraperitoneal injection assay, with death as the read-out. The assay indicates activity in Units or Mouse Units, with 1 U of activity defined as the LD50 dose. However, the manufacturers of the different toxins use slightly different procedures in conducting the assay, and the medical literature indicates that these differences can lead to dramatic differences in the apparent potency. Thus, any clinical differences observed in the number of Units of toxin to achieve specific clinical effects may be partly related to differences in the laboratory assay used to calculated the potency. There are likely to be differences in the affinity of each toxin serotype for its receptor, as well as different receptor numbers per terminal for the different receptor types. These differences may contribute to different clinical characteristics for the different toxin types.

There are also known species differences in sensitivity to the toxins. Guinea pigs appear to be particularly sensitive to botulinum toxin toxicity (compared to some other small animals) but rats seem particularly less sensitive to Type B toxin than to Type A. Therefore, extrapolation of any relative doses for efficacy or toxicity between toxin types based on animal comparisons of effect between toxin types must be done with caution.

Botulinum Toxin Type B

Botulinium toxin Type B is produced by [redacted]

[redacted]

The majority of the clinical development program was conducted with toxin produced at the [redacted]. As clinical development began to auger well for eventual marketing approval, plans were made for moving commercial production to a new facility constructed for the purpose, NeuroBloc Production Facility (NPF) in

California. No clinical experience with toxin produced at NPF was submitted in the initial marketing application. Elan is proposing to market solely the NPF toxin.

The — toxin used in the phase 3 studies, batch A19, had a specific activity of — The three example batches of NPF toxin, designated C90001 to C90003, had specific activities of — to — Batch A19 showed — nicking, with the NPF batches showing— to — nicking. Release specifications for bulk product batches are – to — specific activity, with— to — nicking.

Finished drug product is provided in 3.5 ml glass vials, of liquid toxin solution at a concentration of 5000 U/ml, in dosage amounts of 2500, 5000 and 10000 U (0.5, 1.0, 2.0 ml). The formulation includes 10 mM NaScuccinate, 100 mM NaCl, 0.5 mg/ml HSA, at pH 5.6. The weight of toxin placed into the vial will vary with the specific activity, and may range between — and —

In a small monkey study, IM doses of 1440 U/kg produced systemic, clinically observable muscle weakness in only 1/4 monkeys, doses of 1920 U/kg in 3/4 monkeys, and a dose of 2400 U/kg was fatal in the one monkey tested. Thus, in monkeys 1400 U/kg can be expected to have a significant percentage of animals with systemic weakness, and little more than 50% higher will have fatal effects in many. In a 50 kg person, 1400 U/kg is 70,000 U of Type B toxin.

Cervical Dystonia

[Note to reader: While the general discussion regarding the disease is drawn from multiple sources, including published textbooks, it is most directly from the cervical dystonia discussion in the electronic textbook Neurobase, Gilman S, Goldstein GW, and Waxman SG, eds; Arbor Publishing Corp., 1999 edition]

Cervical dystonia is a syndrome consisting of abnormal head and neck posture with sustained or intermittent movements, and is commonly associated with pain. Previously known as "spasmodic torticollis," cervical dystonia was defined as "an involuntary hyperkinesis involving the muscles of the neck primarily on one side" (Foltz et al 1959). Earlier in this century, the etiology of cervical dystonia was controversial, regarded as psychogenic by some physicians. However, in recent times the organic nature of cervical dystonia has been widely accepted.

Patients with cervical dystonia have involuntary head and neck movements resulting in abnormal postures. The most prominent feature is usually sustained deviation of the head to one side. Terms such as "torticollis," "anterocollis," and "retrocollis" describe the direction of head movement laterally, forwards, and backwards, respectively. There may be lateral flexion of the cervical spine, or horizontal displacement of the head. There is frequent associated asymmetric hypertrophy of neck muscles, the sternocleidomastoid being most commonly involved. Superimposed on the sustained abnormal posture may be fast components in the form of spastic jerks or head tremor, but these are not universally present. Neck pain is a common feature, present in more than half of patients. This pain is often amongst the most troubling aspect to the patient.

Patients may touch certain parts of their head with their hand, and by doing so they may easily bring their head back straight. This phenomenon is known as "sensory trick" and is helpful in establishing the diagnosis of idiopathic cervical dystonia. Associated postural hand tremor is

common, present in about 30%. Typically there should be no contractures, but in patients with a prolonged history of cervical dystonia, there may be fixed deformities in the neck. The abnormal head and neck movements disappear when the patient is sleeping. Swallowing functions may be abnormal, especially in patients with extreme retrocollis.

The etiology of idiopathic cervical dystonia remains unknown. The most accepted theory is that of an abnormality in certain parts of the basal ganglia or brainstem. Since putamenal lesions have been shown to cause contralateral dystonia, the anatomical substrate may be related to this structure or its pathways. However, no definite pathological abnormality has been defined.

The relation with trauma is unclear. Clinically, torticollis occurring shortly after neck injury differs from typical idiopathic cervical dystonia in that there is usually no improvement during and after sleep, and no help by "sensory tricks" (Truong et al 1991).

The prevalence rate has been estimated as approximately 9 in 100,000 (Nutt et al 1988). The overall incidence rate was estimated as 1.2 per 100,000 (Claypool et al 1995). Other sources estimate the US population with CD to be approximately 80,000.

The peak age of onset is from 40 to 49 years, with the majority of patients having onset of the disease between the age of 30 to 55, though it may involve the extremes of ages. A slight female preponderance of approximately 65% women to 35% men has been reported, with various published studies reporting ratios of 1.5 to 1.9 (e.g., Jankovic J., et.al., 1991, Neurology 41:1088-1091; and Chan J., et.al., 1991, Movement Disorders 6:119-126).

The diagnosis is usually made clinically. There is no confirmatory test for cervical dystonia and excluding secondary causes is most important. Spontaneous remission can occur in a minority of patients, usually taking place within the first year of symptoms and with decreasing frequency as the illness becomes more chronic. The majority of patients will have symptoms that usually remain static 5 years after the onset.

Current Management of Cervical Dystonia

There are no currently approved treatments for cervical dystonia in the U.S. Oral medications have generally been disappointing in their effectiveness. High-dose anticholinergic drugs such as trihexyphenidyl have been described to be effective in a small number of patients. Treatment with both muscle relaxants (such as Lorazepam and other benzodiazepines) and spasmolytic agents (such as Lioresal) is a popular combination. Other drugs tried in cervical dystonia include cholinergic, dopaminergic, and antidopaminergic drugs. These may be helpful in individual patients, but the effects may be transient, lasting only a few months, and do not uniformly benefit the broad population of patients.

Botulinum toxin injected intramuscularly into the dystonic neck muscles is currently regarded as the mainstay of relieving symptoms of cervical dystonia. Over the past decade or so many clinical investigators have conducted small studies of varying quality with one of the marketed botulinum toxin Type A products. Frequently, but not uniformly, they have reported favorable results. A National Institute of Health Consensus Development Conference in 1990 issued a

statement that BOTOX was regarded as an accepted therapy for treatment of CD. In the US, only one brand of botulinum toxin, BOTOX, a Type A toxin, is commercially available. BOTOX is not currently labeled for treatment of cervical dystonia, and has significant off-label use in the US for these patients. BOTOX is also approved for marketing in Canada, Europe, and numerous other countries. In many of these countries, but not all, the approved uses include cervical dystonia. A second type A toxin, Dysport (manufacturer previously known as Speywood, now Ipsen) is also marketed in Europe, and is also widely used in Europe for treatment of cervical dystonia. The two toxins are not equivalent nor interchangeable on a unit for unit dose basis. Development of antibodies rendering the toxin ineffective is a concern for long-term usage.

Surgical treatments include thalamotomy, myotomy and rhizotomy, and selective rhizotomy. Varying success has been reported with these techniques, but the results are not uniformly favorable.

Adverse Effects of Botulinum Toxin in Treatment of Cervical Dystonia

The major concerning adverse effect reported in the literature with use of botulinum toxin Type A for treatment of cervical dystonia has been dysphagia. Many investigators have attributed this to spread of toxin locally to pharyngeal muscles adjacent to muscles injected with the toxin. This hypothesis as a source of adverse effects is supported by preclinical studies showing such local spread to adjacent muscles and by observations in treatment of forearm dystonia where EMG localization allowed precise placing of the injections, and specific non-injected but adjacent muscles can be reliably identified and assessed. These adjacent muscles demonstrated loss of strength in 63% of subjects (n=40, Ross MH, et.al., 1997, Muscle and Nerve, 20:593-598).

Most subjects in published studies have had only mild to moderate severity dysphagia. Nonetheless, some subjects do discontinue repeat injection sessions due to the dysphagia. However, occasional subjects in studies have needed nasogastric feeding tubes until improvement of swallowing. This adverse event has consistently resolved over time (Stell R, et.al. 1988, JNNP 51:920-3). Other studies have reported subjects who required iv fluid therapy for severe dysphagia, with aspiration changes observed on chest xray.

Anderson reported use of Dysport in 107 subjects, with repeat injections for a total of 510 treatment sessions, of which 2% of treatments had severe dysphagia. Two treatment sessions lead to hospitalization for assisted hydration (duration unstated) and two other sessions lead to substantial weight loss. One subject developed aspiration pneumonia. No deaths were reported. The dysphagia was believed to be dose related, especially the dose injected into the SCM muscle. Anderson also reported an event of leg weakness in a subject with a prior history of polio, suggesting hematogenous spread also and sensitivity of muscles in this patient. (Anderson TJ, et.al., 1992, J. Royal Soc Med. 85:524529)

There have been rare reports of severe dysphagia in subjects with co-existing known as well as unrecognized neuromuscular disease. Tuite and Lang (1996, Neurology 46:846), report 2 subjects with known Machado Joseph Disease who were treated for accompanying cervical dystonia. These subjects developed severe dysphagia which persisted for months, after receiving 320 or 250 U BOTOX. Gastrostomy was required for 6 months in both patients. Emmerson

(1994, Mov Disord 9:367) reported a subject with myasthenia gravis treated with botulinum toxin, who also developed severe dysphagia.

Erbguth et.al. (1993, JNNP 56:1235-6) report on a patient treated for blepherospasm with Dysport (8ng) who experienced marked generalized weakness as a result which lead to a diagnosis of paraneoplastic Lambert Eaton Myasthenic syndrome, which had not been clinically evident prior to toxin injection. The authors emphasize that patients with underlying neuromuscular disease are at increased risk of generalized muscle weakness from even low doses of botulinum toxin injected for local effect. Borodic (1998, Lancet 352:1832) reports a similar case of uncovering a diagnosis of myasthenia gravis in a patient where the dysphagia was severe enough to require a gastrostomy for nutrition.

Bakheit et.al. (1997, JNNP 62:198) report a multiple sclerosis patient who received 250 U Dysport which lead to widespread weakness and a patient with multisystem atrophy and existing pharyngeal dysfunction who had been receiving without problems 750 U Dysport for 5 years of repeated injection, who then experienced severe generalized weakness after a regular treatment session. Mezaki et.al. (1996 Neurology 46:845) report on an ALS patient who received 300 U of a Japanese marketed botulinum type A product and had a dramatic increase in weakness.

Taken together, these reports emphasize that patients with neuromuscular disorders may be markedly more systemically sensitive to the generalized effects of local IM toxin injection than patients without such disorders.

Assessment Scales for Cervical Dystonia

Several clinical investigators have developed their own evaluation procedures for CD, and published study results employing these. However, widespread use in the medical literature has been limited to two multicomponent assessment scales and to subjective "global assessment" scales.

Tsui and colleagues at U.British Columbia, Canada invented the tool now known as the Tsui Scale. This scale has 4 questions relating to the aspects of severity of the dystonia which are given point values and then the first two are multiplied together. Both the two remaining questions are then added to this product for the final total. The maximum score is 25, and this scale is reported to have a good interrater correlation for scores between 5 and 20 (Tsui JKC, 1986, Lancet ii:245-247). The Tsui scale does not incorporate a measure of pain. As pain can be a prominent feature of CD in many patients, and has been reported to be an important component of the beneficial effects of botulinum toxin treatments, the Tsui Scale alone is regarded as inadequate to fully assess CD outcome. Clinical studies employing the Tsui Scale generally also employ a simple pain severity question to supplement the scale. The Tsui Scale appears to have been the most widely used scale in the medical literature. Baseline median score in CD studies is approximately 11 to 13.

Consky, Lang and coworkers in Toronto, Canada, developed the Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS) in the late 1980's and early 1990's. This scale employs multiple questions in each of three subscales, rating severity of dystonia, patient perceived disability resulting from the dystonia, and pain from the dystonia (Consky ES and Lang AE,

1994, Clinical Assessments of Patients with Cervical Dystonia, in: Jankovic J and Hallet M, eds. Therapy with Botulinum Toxin. NY: Dekker, p 224-226). Each question contributes from 0 to a maximum of 2 to 5 points, which are then combined in a complex manner (some questions averaged, some question's scores are multiplied together) to give a scale that ranges from 0 (no symptoms) to 87 (maximum severity, disability and pain). The Tsui Scale and the TWSTRS-Disability Subscale have considerable overlap in content of questions. An earlier form of the assessment was reportedly evaluated for reliability and found to have fair to good interobserver agreement (Consky ES, et.al., 1990, Neurology, 40 (Suppl 1):445) but this abstract report was never published with full details and had assessed an early form of the scale at least partly different from the form commonly in use at present. Multiple studies employing this scale have been published. This scale incorporates a pain assessment, thus not requiring a separate pain evaluation method.

In addition, many studies also use a patient or investigator global assessment scale. These evaluations are often different in detail, but generally similar in concept. The rater is asked to select a score from a range (e.g. 0 to 10, 0 to 3, or draw a mark on a 100 mm line visual analog scale) that provides an overall impression of the status of the patient at that time.

A study seeking to compare these scales has been conducted (Tarsy, D., 1997, Movement Disorders 12:100-102). There were 76 subjects with CD assessed with TWSTRS, Tsui, and Physician Global rating scale (a simple 0-3 scale), before and after treatment with BOTOX. Tarsy found moderate correlations of TWSTRS-Total with the Global assessment, and of TWSTRS-Severity with Tsui scales, with somewhat lesser correlations of Tsui Scale with Global Assessment, and TWSTRS-Total with Tsui. Correlation of any of the subscales of TWSTRS with another TWSTRS subscale was low, thus supporting the intent that the subscales are assessing different aspects of the disorder and it's effect upon patients.

Thus, the medical field appears to have accepted both Tsui and TWSTRS Scales as valid assessment tools for CD. The TWSTRS is more comprehensive, but also a more complex composite and more difficult to interpret, especially for small changes in score. The Tsui Scale is simpler to interpret, but cannot suffice alone as pain is not assessed within the Tsui Scale. No clear basis for determining a minimal meaningful change in score has been established for either of these scales.

PHASE 3 STUDIES AN072-301 and AN072-302

OVERVIEW

Studies AN072-301 and AN072-302 (hereafter referred to as simply 301 and 302) are the two primary sources of controlled safety and efficacy data for the use of BotTx-B for the treatment of cervical dystonia. These were two separate studies, but conducted concurrently, at many of the same study sites, and employed protocol procedures that were largely identical. They will be presented in parallel in this review document due to their similarity.

In overview of design, both studies enrolled patients with well established cervical dystonia of at least moderate severity, and in whom it was known from prior clinical management had been responsive (obtained benefit from) to BOTOX (botulinum toxin type A) in off-label use. The subjects were randomized in a double blind manner to a single treatment cycle with either BotTx-B or placebo, with follow-up for the succeeding 4 months. The only significant differences between the two studies were that Study 301 randomized to three treatment groups (placebo, 5000U, or 10,000U) while Study 302 randomized to two treatment groups (placebo or 10,000U) and that Study 301 enrolled subjects who's CD symptoms were still responsive BOTOX while Study 302 enrolled subjects for whom BOTOX had subsequent ceased to provide benefit, and were regarded as unresponsive to BOTOX, both in CD symptoms and in paralysis of what ever muscles were injected.

Study 301 Title: A double-blind, placebo controlled, single dose, safety and efficacy study of BotB (Botulinum Toxin type B) in patients with cervical dystonia.

This protocol was initially finalized February 1997, and then subsequently amended twice, April 1977 and May 1977. Both of these amendments occurred prior to study initiation, and the protocol as conducted is presented in this review. The final analytic plan was submitted in March 1998, after completion of the study, but prior to unblinding of the study results. These analytic plan changes are also incorporated in this study presentation.

Study 302 Title: A double blind, placebo controlled, single dose, safety and efficacy study of BotB (Botulinum toxin type B) in Type-A resistant patients with cervical dystonia.

This protocol was initially finalized in February 1997, amended in April and twice in May 1997, prior to study initiation. The protocol was conducted as described in this review. The final analytic plan was revised in March 1998 (simultaneously with that of Study 301), after completion of the study but prior to unblinding of study results. These analytic plan changes are also incorporated into the study description.

CLINICAL STUDY DESIGN

Objectives

Study 301: To evaluate the safety and efficacy of BotTx-B in subjects with cervical dystonia.

Study 302: To evaluate the safety and efficacy of BotTx-B in subjects with cervical dystonia, and to demonstrate that BotTx-B is efficacious in subjects unresponsive to Type-A toxin.

General Design Structure

Both of these studies were multicenter, entirely outpatient double blind, placebo controlled, parallel group studies. Eligible subjects receive a single treatment session with the assigned study drug. Subjects then return to clinic at specified intervals for evaluations for up to 4 months.

For study 301, 108 subjects were planned for enrollment, divided into 3 parallel groups (36 per group). For Study 302 80 subjects were planned for enrollment, divided into 2 groups of 40 each.

All investigators were to obtain local IRB approval and obtain written informed consent from subjects prior to performing any study-related activities.

Eligibility Criteria

Inclusion criteria

1) History of CD for at least 1 year, affecting at least 2 of the following muscles:
   - Leavator sacapulae
   - Scalenus medius & Scalenus anterior
   - Semisinalis capitus
   - Splenius capitus
   - Sternocleidomastoid
   - Trapezius
2) CD of at least moderate severity on the TWSTRS scale:
   - Total >= 20 points
   - Severity >= 10 points
   - Disability >= 3 points
   - Pain >= 1 point
3) Known previous history of meaningful clinical response of CD symptoms to BOTOX
4) Male or Female, >= 18 yo
5) Weight >= 46 kg

6) Known current status of BOTOX responsiveness

Study 301: Last injection cycle with BOTOX continued to provide worthwhile benefit.

Study 302: Now known to be unresponsive to Type-A toxin, as shown by:

(A) (1) Failed to respond to prior 2 treatment sessions with BOTOX, and One of the two sessions at higher dose than previously efficacious

or

(2) Failed to respond to last treatment session with BOTOX and either

(I) subject has a known positive mouse neutralization test for Ab

or

(ii) single muscle injection at higher dose of Type A toxin was ineffective

And

(2) F-TAT demonstrates lack of activity of Type A toxin

F-TAT (frontalis- type A test) consists of injection of two 7.5 U doses of BOTOX into the R. frontalis muscle, positive test indicated by subjects ability to continue to wrinkle the forehead on the right at 2 weeks later.

Exclusion Criteria

1) Receipt of BOTOX injection within prior 4 months.
2) Prior receipt of BotTx-B
3) Neck contractures or cervical spine disease causing decreased neck ROM
4) Pure retrocollis or anterocollis
5) Women who are pregnant, breast feeding, not using adequate contraception
6) Irregular use of narcotics, muscle relaxants, or benzodiazepines for treatment of CD symptoms. Regular use of these medications permitted if planned to be sustained during study.
7) Other acute or chronic medical conditions that preclude administration of BotTx-B.
8) History of myotomy or denervation surgery in neck or shoulder region.
9) Tetanus toxoid receipt within prior 4 months.
10) Use of another investigational device or drug within prior 30 days
11) History of known neuromuscular disorder, or other significant organ disease

Study Treatments and Concomitant Medications

Study Treatments

Study 301: Subjects were randomized to 3 groups, placebo, 5000U or 10000U.
Study 301: Subjects were randomized to 2 groups: placebo or 10000U.

Study drug was supplied as a solution in a 3.5ml vial containing 5000 U of BotTx-B in 1ml of 0.05% human albumin, 0.01M succinate/NaCl buffer at ph 5.5 (0.5mg HSA, 1.6mg NaSuccinate, 5.8mg NaCl per ml). Placebo vials contained 1ml of the buffer without any toxin. Vials were to be stored in a 2-8 °C refrigerator until use (freezing prohibited). Lots used for these two studies

were from Batch A19 which had been produced in the — facility. No product produced in the NPF facility were used.

The vials were packaged into patient vial boxes, each containing two vials. The placebo and toxin vials were appropriately packaged together to make up placebo, 5000U or 10000U total doses when the contents of the two vials were combined prior to subject injection. The entire 2 ml were to be injected into each subject. Vials and boxes had white labels for Study 301, blue for Study 302 to avoid confusion at study sites conducting both studies.

Investigators were instructed to use their own judgement to select 2 to 4 muscles for injection, and to allocate the dose between those muscles as directed by individual judgement. Muscles eligible for injection were the same as those permitted to qualify the subject for inclusion in the study. Intramuscular injection of the study agent was performed in the selected muscles, using 1 to 5 sites per muscle, at investigator discretion.

Investigators were selected to be those with prior established experience in treatment patients for CD with BOTOX in an off-label manner. EMG guidance of the injections was at the discretion of the investigator.

Concomitant Medications

All pre-existing use medications (used at entry), both prescription and over the counter, were intended to be continued in the same regimen during the study. New medications were to be avoided if possible. Irregular use of narcotics, muscle relaxants, benzodiazepines were to be avoided if possible during the study, but if they could not be avoided, were to be prohibited for at least 24 hours prior to each outcome evaluation during the study.

Blinding

The patient vial boxes and vials were identified by a Patient number on the blinded study labels. At the time of use, label portions were attached to CRFs to identify the drug administered. An emergency unblinding disclosure panel (employing scratch-off material to maintain blinding) was included with the patient vial box and attached to the CRF.

To assist in maintaining the blind, two investigators were used at each site, a Principal and an Administrative investigator. Principal investigators conducted the screening, the treatment, and the outcome efficacy evaluations. Administrative investigators conducted all other subject interactions and evaluations.

Randomization

Patient study boxes sent to each site were sequentially numbered, and sites were instructed to use the boxes in sequence as subjects were enrolled. The study vial box number then becomes the Patient Identification Number. Study kits were uniquely numbered, so that only one kit #3 existed within the study, etc. Thus, a contiguous sequence of kits, not starting at # 1, were sent to each site. Randomization was performed by blocks within study site.

Subject Evaluations

Medical history, physical exam, neurological specific history and exams were performed at the start of study participation and at the end. Vital signs, clinical laboratory assessments, adverse event monitoring and medication usage recording were performed throughout the study period at subject visits to the study site. A phone call to the subject 1 week after study injections for adverse event monitory was also performed. Serum samples for antibody testing were obtained at baseline, week 4 and the end of study participation. In addition, study assessments at the scheduled study visits included the following:

Screening Visit (up to 21 days prior to day 0)
- Patient Global VAS
- Patient Pain VAS
- TWSTRS evaluation

Day 0 (Baseline; day of treatment, administered after evaluations)
- Patient Global VAS
- Patient Pain VAS
- TWSTRS assessment
- Investigator Global Assessment Notes - taken by PI for reference in completing Inv.Global Assessment at subsequent visits

Week 2, 4, 8, 12, 16
- Patient Global VAS
- Patient Pain VAS
- TWSTRS Assessment
- Investigator Global VAS

Timing of Subject Visits

Due to potential disease variability during the day, the study intended that each patient be evaluated at the same of day for each of the study visits.

Week 2 and 4 visits were to be within 3 days of the exact scheduled timing; week 8, 12, 16 visits were to be within 1 week of the intended schedule.

Description of Evaluation Tools

TWSTRS: Toronto Western Spasmodic Torticollis Rating Scale
- A composite of a variety of features of the disease and effects on patient life, with 0 as a no-symptoms score, and 87 the extreme maximum worst possible score.
- The composite is broken into three main category subscales: Severity, Disability and Pain.

Severity Subscale

5 questions on amplitude of the symptom, each a maximum of 2 - 4 points
1 question on duration of maximal or submaximal symptoms, up to 10 points
4 Additional questions on relief from sensory tricks, amount of shoulder displacement, amount of active ROM, and length of time subject can maintain head in neutral position.
Points from each question summed, maximum severity is 35 points.

Disability Subscale

7 questions, each 0 to 4 or 5 points, on employment activity, ADL in general, driving, reading, television watching, activities outside of home, and amount of embarrassment limitation
Points from each item summed, maximum is 32 points

Pain Subscale

Pain Severity rated for best, worst, usual, and then composite of the 3 taken, for up to 10 pts.
Duration of pain, and Amount of disability that results from just the pain are rated.
Severity, duration and disability contribution summed, for maximum of 20 pts.

The scale maximum of 87 points should not be regarded as a clinically observed score. Rather the scale is a variety of CD aspects, combined in somewhat arbitrary manners, which accommodate the fact that different patients have different aspects as their most prominent and limiting characteristic. The scale is intended to capture the disease status for all patients, irrespective of which is their most prominent.

Patient Global Assessment of Change VAS

A horizontal line, with only end bars (no central marker), of 100 mm length, left end labeled "much worse", right end labeled "much better". The printed instructions to patient are to draw a vertical line to indicate how they felt at the visit compared to prior to study treatment.

For screening and day 0 visits, in order to accustom the subject to completion of the tool and to assess the variability in the questionnaire results the VAS was also completed. At these visits, however, the phrase directing the assessment was compared to 1 week ago.

The score was the number of mm from the left end, so that 0 was the worst score, 100 best. While the central point of the scale was not marked, the implicit interpretation is that a score of 50 means no difference from prior to study treatment. This evaluation tool implicitly caries the meaning of change from baseline for all post-treatment scores.

Patient Pain VAS

A horizontal line of 100 mm length, left end labeled "worst ever pain", right end "no pain", with only end vertical bars. Instructions to subject are to draw a vertical line to

indicate the intensity of the pain from CD over the week preceding the visit. The instructions do not indicate whether least, worst, or typical pain should be rated. This evaluation is one of absolute amount of severity over the immediately preceding week. Interpretation of individual patient scores requires reference to the subject's baseline Pain VAS score.

Investigator Global Assessment VAS

A horizontal line with only end bars of 100 mm length, marked on left end "much worse" and right end "much better", identical to that used for the Patient Global Assessment. Instructions to investigator are to draw a vertical line to show how investigator believes the subject is at time of visit compared to prior to study treatment. Instructions to the Investigator were to include consideration of factors such as daily activities, psychological status, pain, muscle tone, general functional status.

Endpoints and Planned Analyses

Efficacy Endpoints

Primary: TWSTRS Total score at Week 4
(Pairwise comparison of placebo to 10000 U group, for both studies)

Secondary: Patient Global Assessment of Change at Week 4
(Pairwise comparison of placebo vs 10000 U group, for both studies)

Tertiary: Investigator Global Assessment at Week 4
TWSTRS-Total score at Week 8, Week 12
Percentage of Responders
Patient Pain VAS scores at Week 4

Additional Endpoints: TWSTRS Total scores at other visits,
TWSTRS Subscale scores at visits
Study-301: All analyses of the 5000 U group results.

Analytic Plan

There was no Interim Efficacy Analysis planned for these studies.

Two datasets were prepared, an Intent to Treat and a Modified Intent to Treat. The ITT dataset was prospectively declared the primary basis for assessment of study results. The ITT dataset has a value for all subjects at all visits for all evaluations. Missing data for efficacy variables were filled by LOCF. If a subject has no prior data for a variable at missing visit, a value for that subject is imputed at the first missing visit by using the mean of all subjects in the study with values at the visit. This imputed value is then used for subsequent LOCF. An exception is for responder/non-responder analysis, where if data is missing at week 4, the subject is deemed a non-responder.

The Modified-ITT dataset includes subjects who received study treatment and had at least one outcome evaluation post baseline. LOCF was applied only to missing values for the week 4 outcome if the week 2 visit was available. Missing values at all other visits were left missing, and the subject excluded from the analysis for that variable/timepoint.

Significance was determined by $p<0.05$ for all tests. Interaction terms were deemed significant if $p<0.10$.

The primary endpoint of TWSTRS Total score was analyzed with ANCOVA, with covariates of Baseline TWSTRS score, Center, and Baseline x Treatment and Center x Treatment interactions. Interactions will be dropped if their p-value in the model is >0.10. Adjusted means will be calculated for centers and treatment group.

The secondary endpoints of Global Assessments analyzed with ANOVA (treatment, center and interaction included), TWSTRS at Week 8 and 12 analyzed with ANCOVA.

Percentage of Responders analyzed with Cochran-Mantel-Haenszel, Patient Pain VAS analyzed with ANCOVA (treatment, center, and baseline Pain score, with interactions included).

Elan (Athena at that time) had prospectively declared that the responder analysis was based on a definition of a responder as 20% improvement from baseline in TWSTRS Total score at the Week 4 visit.

Comment:

This criterion had not obtained concurrence from CBER despite extensive discussions with Athena prior to unblinding the study. Despite a panel of experienced investigators brought together by Athena, no adequate basis for selecting or justifying the 20% change from baseline criterion was presented. Partly in recognition of this disagreement, Athena designated this analysis as a tertiary endpoint, rather than making it a secondary endpoint.

The primary analysis of the TWSTRS scores was specified in the analytic plan as being conducted on the absolute TWSTRS score. Change from baseline in TWSTRS was also specified for analysis, but absolute score is stated as the primary analysis of the outcome.

Comment:

However, an ANCOVA adjusted with baseline TWSTRS included is the primary analytic method. For the calculation of a test statistic and p-value, this is highly similar to subtraction of baseline score. Thus, it is mostly change from baseline that is being compared rather than simply the outcome scores, and the appropriate clinical outcome to examine for assessment of the amount of benefit is change from baseline.

Planned Sample Size

For Study 301, sample size was based upon a 2-sided F test at 0.05 level of an overall treatment effect (amongst the 3 groups). A study with 27 subjects per group provides 90% power to detect a treatment effect of 12 points between at least 2 of the treatment groups, when the standard deviation is also 12 points. The sample size was arbitrarily increased to 36 per group to increase the database size for safety assessments, and the ability to demonstrate a difference between the placebo and 5000U group.

For Study 302, the sample size was based upon a 2-sided F test at 0.05 level, with 23 subjects per group providing 90% power to detect a treatment effect of 12 points when standard deviation is 12 points. Sample size was arbitrarily increased to 40 subjects per group to provide increased safety data and to analyze secondary endpoints.

Protocol Modifications

There were no significant modifications to the study protocols after the start of the study other than the analytic plan. The analytic plans underwent significant modification during the study and was not finalized until after the data collection had been entirely completed. The analytic plan was finalized prior to unblinding the study and commencing analysis.

STUDY PERFORMANCE AND SUBJECT DISPOSITION

Enrollment and Subject Disposition

The two studies were conducted largely concurrently, and substantially at the same sites. These studies were both conducted from May 1997 to February 1998.

Study 301

A total of 109 subjects were enrolled, and randomized as 36 to placebo, 36 to 5000 U, 37 to 10000 U. There was one error in randomization. At site 5, treatment kit 13 is reported as leaking as the vials were being prepared for use on the first subject at the site, and was discarded. Kit 14 was used in it's place for that subject. Kit 174 was provided to Site 5 and intended for use on the next enrolled patient (who initially would have been subject #14), but was not. Kit 15 was used for that next subject erroneously. Kit 174 was subsequently used for the following subject, who originally would have received Kit 15 at that site. The actual treatment received of these three subjects was 5000 U - 10000 U - placebo. The prospectively intended treatment assignments had been placebo - 5000 U - 10000 U. These three subjects were analyzed in the dataset according the actual treatment received. These subjects are identified in the dataset by the Kit # as received. No subject #13 exists in the dataset.

There were 4 subjects who discontinued participation in the study early. Two (one placebo, one 5000 U subject) for perceived lack of efficacy, one for a change in workplace (a placebo subject), and one for a serious AE (10000 U subject). This subject had a MI, which lead to undergoing CABG procedure and died due to cardiac arrest several days post CABG. These subjects exited from the study after study visits at weeks 8, 12, 0, 4 respectively (i.e., only one of these subjects did not have actual visit data for the primary endpoint at week 4). There were 105 subjects who completed all evaluations in the study.

Study 302

A total of 77 subjects were enrolled and randomized 38 to placebo, 39 to 10000 U of toxin. There were no unblinding events, and no errors in randomization. There were 7 patients who were screened for the study but in spite of apparent Type-A resistance therapeutically, failed to show resistance on the F-TAT.

One subject discontinued the study early, a placebo subject due to adverse event of headache and neck pain due to CD. This subject had AE onset several days after study treatment, and did not return for any study follow-up evaluations.

Time in study

Most subjects completed a full 16 weeks participation in both of the studies.

Protocol Deviations

Study 301
There were no major protocol deviations. Errors in treatment assignment were as noted above. Two subjects signed the wrong consent form version, and one subject had a history of neck surgery, but this patient was approved for enrollment by Athena after discussion with the investigator. There were several instances of laboratory values (CBC, chemistries) being reviewed by the Principal Investigator rather than the Administrative Investigator, and 25 instances of a scheduled evaluation not being performed even though the subject had shown up for clinic visit (mostly vital signs and some Global Assessments)

Study 302
There were 9 subjects with minor errors of eligibility. These include a history of surgery for CD, history of neuromuscular disease, incorrect type of antibody testing for antibody to Type A toxin, and 4 subjects who had their F-TAT read at a time other than the specified 14 days.

A comparison of the actual timing of study visits compared to the scheduled timing was not assessed by the Applicant.

Study Sites and Site Enrollment

There were 12 different sites participating in at least one of these two studies. Study 301 had 9 sites, Study 302 had 7 sites, with only 4 sites overlapping between the two studies. Because blocking within sites had been used, with block size of 3 for Study 301 and blocks of 2 for Study 302, there was good balance in group sizes within each center.

Table 1: Study Sites and Enrollment by Group

Site Code #	Princ. Investigator	Study 301			Study 302	
		Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
4					3	4
5		4	4	4	5	5
8		5	5	4		
9		4	4	4	5	5
12		3	2	3		
17		2	2	2	8	8
18					5	5
19		5	5	5	6	6
20		4	4	4		
22		6	6	7		
23		3	4	4		
24					5	6

Blank cells indicate no subjects enrolled at that site to that study

Treatment Administered

Elan did not submit a summary of the treatment characteristics.
Comment:

A summary of treatment, including number of muscles per subject, utilization rates for specific muscles, range of doses per muscle, etc., will be requested.

Bioresearch Monitoring Inspections

Bioresearch Monitoring initiated inspections at 4 study sites, of which 3 EIRs had been returned as of the date of writing this review. While some inconsistencies had been discovered, none were major enough to call the overall integrity of the studies into question. CBER Bioresearch Monitoring reports that the deviations were not substantitive.

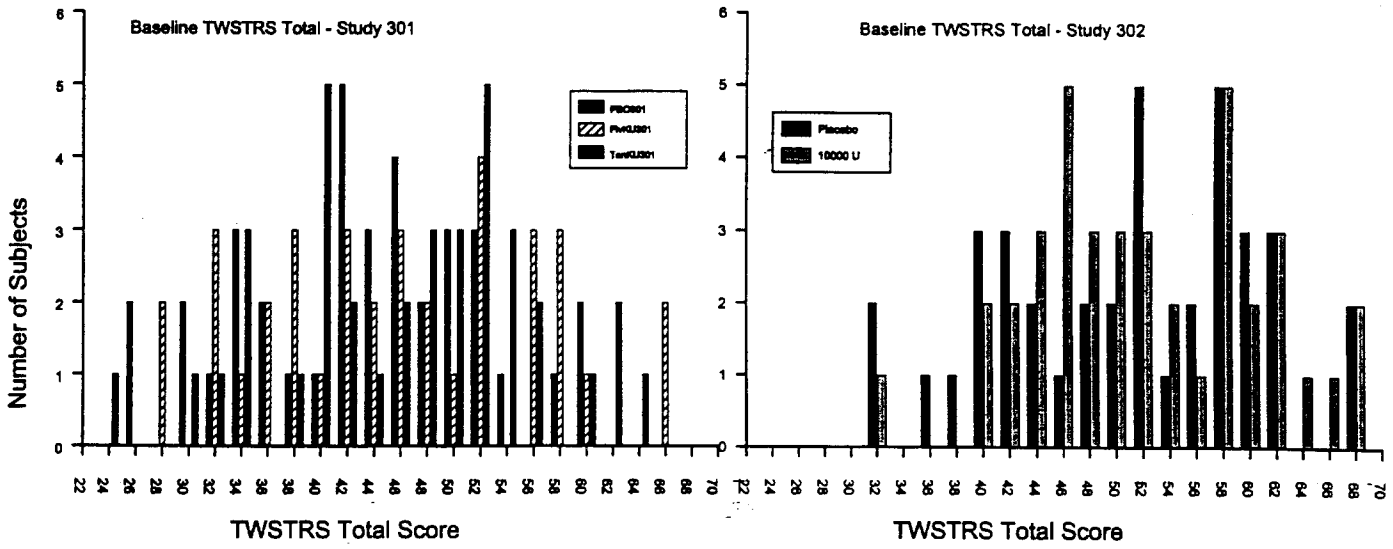
Study Population Characteristics

The study population characteristics in age and sex distribution are similar to that expected from the medical literature. Baseline disease severity was moderate in these subjects, as expected from the eligibility criteria that excluded more mildly affected subjects. There was good comparability between the treatment groups within a study for the baseline TWSTRS severity, but less so for the baseline Patient Pain VAS.

Table 2: Demographic and Baseline Characteristics						
Parameter		Study 301			Study 302	
		Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Age	(mean, yrs)	54	58	56	53	57
Sex	% female	58	50	76	68	69
	% male	42	50	24	32	31
Race	% White	89	97	89	100	100
	%black	3	0	3	0	0
	% hispanic	6	0	8	0	0
	% other	3	3	0	0	0
Height	(cm, mean)	170	169	167	170	167
Weight	(kg, mean)	74	77	74	75	75
Baseline TWSTRS Total		43.6	46.4	46.9	51.2	52.8
Baseline Patient Pain VAS		43.6	40.8	35.1	33.6	41.4

Histograms of the Baseline TWSTRS score suggest the shape of the distribution of scores within each group were also similar between the groups within each study. While scores as low as 20 were permitted, no subjects in either study had a score that low. Baseline scores were broadly distributed in the moderate to severe range, with a maximum score of 68, in Study 302.

FIGURE 1



EFFICACY RESULTS: PRIMARY EFFICACY ENDPOINT AND SUBSCALES

TWSTRS Scores and Changes from Baseline

The primary Efficacy Endpoint was the TWSTRS Total Score at Week 4, compared pairwise for the placebo vs. 10000 U groups for both studies. This table appears to suggest a modest improvement in TWSTRS at week 4 for all toxin groups compared to the placebo group. However, due to the non-identical baseline TWSTRS means, the week 4 outcome is difficult to interpret in isolation.

Table 3 TWSTRS Total Scores

Time Point	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Baseline	43.6	46.4	46.9	51.2	52.8
Week 2	40.2	37.3	36.6	49.1	42.5
Week 4	39.3	37.1	35.2	49.2	41.8
Week 8	41.3	39.4	38.5	49.6	44.1
Week 12	42.2	42.9	42.8	50.5	46.8
Week 16	43.3	45.5	48.2	50.9	49.6

Means without any ANCOVA adjustments shown

In prospective recognition of this, the analytic method was not simply a comparison of week 4 scores, but a comparison using ANCOVA with baseline TWSTRS score included. The primary endpoints of the Week 4 TWSTRS in the two studies were both statistically significantly different in the toxin groups vs placebo. The p-values from this analysis are shown in the following table, which also provides a secondary analysis that more clearly indicates the treatment associated effect, the change from baseline in the TWSTRS Total score. For this and all change from baseline analyses in this report, change from baseline has been calculated so as to give a positive value for changes that describe improvement of the clinical status. For the TWSTRS scores, this implies change is calculated as Baseline Score - Outcome Score, rather the more common automatic approach of outcome - baseline. For Study 301, the 10000 U vs placebo comparison was stated as the primary endpoint, and indicates efficacy with that toxin dose. However, the study results also support efficacy with the 5000 U dose in this study.

Table 4 TWSTRS Total - Change from Baseline Scores (as Improvement)					
Time Point	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Baseline Score	43.6	46.4	46.9	51.2	52.8
Week 2	3.4	9.1	10.4	2.2	10.4
Week 4	4.3	9.3 *(0.012)*	11.7 *(0.0004)*	2.0	11.1 *(0.0001)*
Week 8	2.3	7.0 *(0.019)*	8.4 *(0.003)*	1.7	8.8 *(0.0002)*
Week 12	1.4	3.5 *(ns)*	4.1 *(ns)*	0.8	6.0 *(0.013)*
Week 16	0.3	0.9 *(ns)*	-1.3 *(ns)*	0.3	3.2 *(ns)*

Change calculated as Baseline - Outcome; Positive => improvement; simple means with adjustment
P-values shown in () for pairwise comparison to placebo from ANCOVA

For the week 4 ANCOVA analysis, Study 302 had significant interactions of treatment assignment with baseline TWSTRS score and study center. Final model was nonetheless calculated without the interaction terms.

The estimated treatment effects were summarized with point estimates and confidence intervals on the change from baseline in TWSTRS Total score:

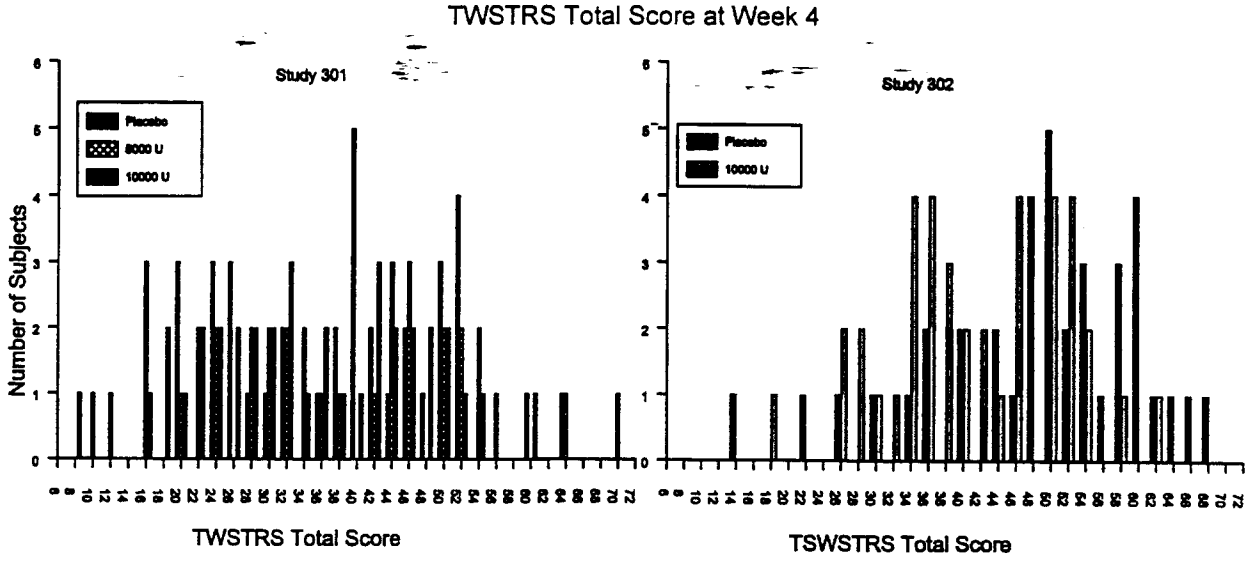
Table 5: Estimated Treatment Effect on TWSTRS									
Time Point	Study 301						Study 302		
	5000U vs Placebo			10000U vs Placebo			10000U vs Placebo		
	est Tx effect	95% CI	p value	est Tx effect	95% CI	p value	est Tx effect	95% CI	p value
Week 4	5.0	8.9, 1.2	0.012	7.2	11.1, 3.3	0.0004	8.7	12.2, 5.2	0.0001
Week 8	4.8	8.8, 0.8	0.019	6.1	10.1, 2.1	0.003	6.8	10.2, 3.4	0.0002
Week 12	2.0	6.0, -2.1	0.33	2.5	6.5, -1.6	0.22	4.8	8.5, 1.0	0.013
Week 16	0.4	4.6, -3.8	0.86	-2.0	2.2, -6.2	0.35	2.8	5.9, -0.4	0.08

Study 301 estimates calculated with adjusted least sq. means
Study 302 estimates calculated with reduced analyses of covariance
Effect shown as decrease in score; positive => improvement

Comment:

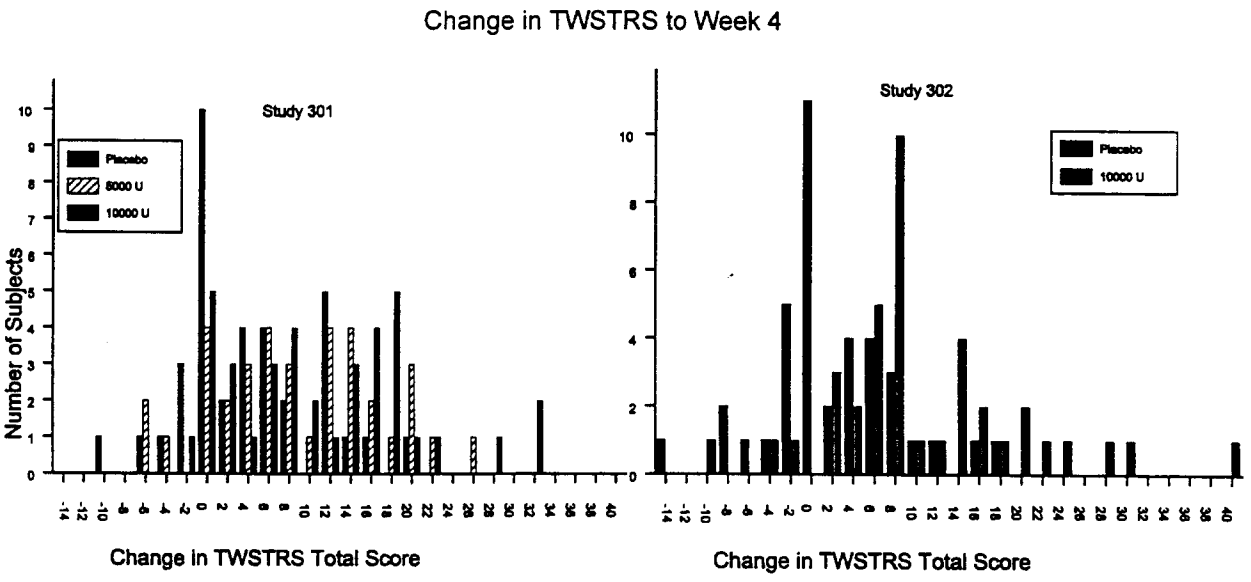
The distribution of the week 4 scores, shown in Figure X below, appear consistent with this analysis, indicating some differences between the treatment groups as well. No anomalous shifts in the distribution not revealed by simple means were evident. Also apparent is that this treatment is far from completely benefiting the subjects. No subject achieves a score of 0 (no symptoms) and almost all subjects remain with a score above 20, which was the moderate level of severity that served as the entry criterion.

FIGURE 2



The distribution of change is TWSTRS at week 4 is more clear in the separation of the treatment group distributions. These histograms also serve to illustrate the variability of the disease. Even the placebo subjects show a range of changes in score at week 4. The placebo distribution is centered about a change of a small degree of improvement, perhaps suggesting a placebo effect and there are a substantial portion of the subjects that have changes of either substantial improvement or some degree of worsening.

FIGURE 3

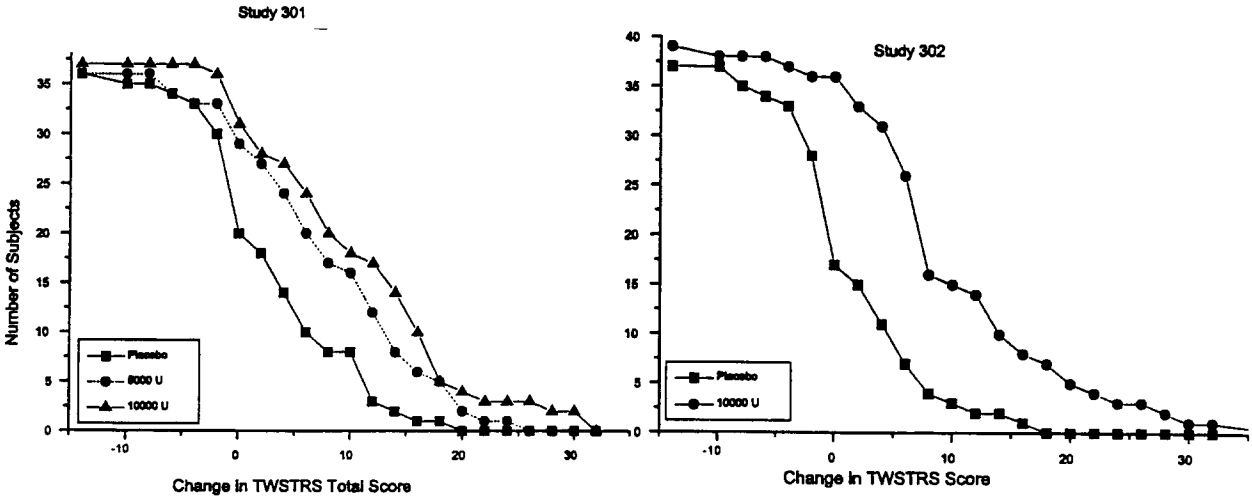


However, the histograms remain difficult to examine to assess the overall amount of benefit derived from the treatments. For this purpose, cumulative distributions are more clear. Curves showing the fraction of subjects who have achieved score changes equal to or greater than specific levels better illustrate the treatment effects upon the group. Figure X, below, shows that in both studies, there was clear separation between the toxin and placebo groups in number of subjects who achieved amounts of improvement. The toxin

groups are shifted to greater numbers of subjects at larger amounts of improvement from baseline compared to the placebo group. This is true for both dose levels in Study 301 as well as the single toxin group in Study 302. The curves for the toxin and placebo groups are generally parallel to each other, suggesting that the benefits were broadly distributed amongst all subjects in the treatment groups, and not confined to a small subset. No subjects receive massive amounts of benefit, but most appear to receive some. Of particular note, however is that in Study 301, the 5000 U group received nearly the same amount of shift as the 10000 U group. There is relatively little evidence of additional benefit with the increased dose at week 4.

FIGURE 4

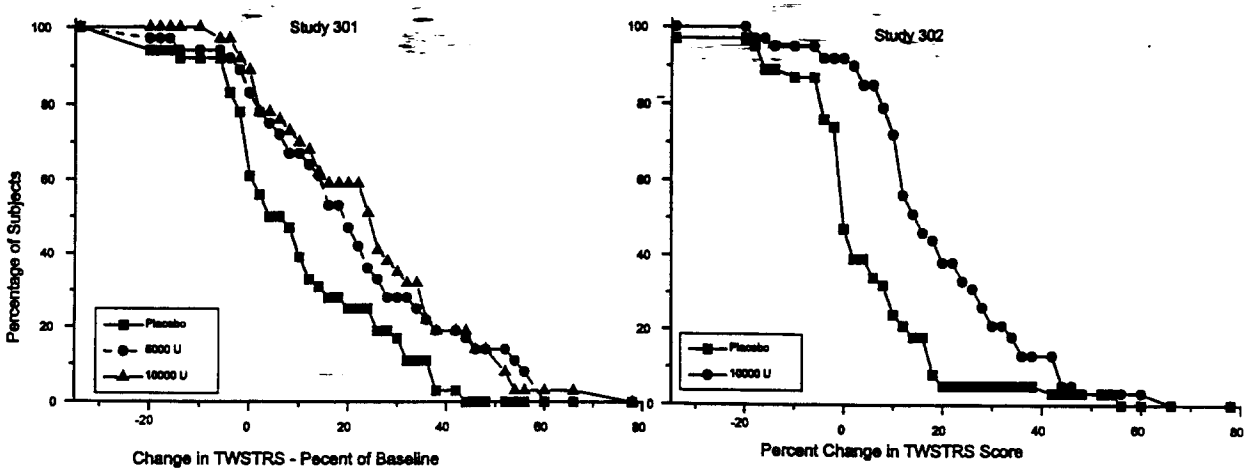
Subjects with Changes in TWSTRS Scores:
Number of Subjects Achieving At Least Specific Amounts of Change in Points



The TWSTRS is not a uniform interval scale, and changes of 3 or 5 or 9 points are not necessarily of equivalent meaning to subjects of different baseline scores. Since subjects enter the study with a broad range of baseline scores, assessing the change as a fraction of the baseline severity has been proposed as another approach to understanding the utility of treatments in cervical dystonia. The cumulative subjects with changes in TWSTRS is shown in Figure X as percent of subjects with at least specific percentage improvements. These curves again indicate that the benefits broadly distributed to subjects receiving the toxin, and that there was little to no greater efficacy in the subjects receiving 10000 U than those receiving 5000 U.

FIGURE 5

Changes in TWSTRS as Percent of Baseline:
Percent of Subjects Achieving At Least Specified Amount of Change



Effect of Missing Data on Primary Endpoint

Comment:

As noted above, there were 3 subjects in Study 301 and one in Study 302 for whom Week 4 visits were not obtained. In study 301 these were distributed one in each treatment group. The data imputed for the ITT analysis was of no change from baseline for the subjects in the placebo and 10000 U groups, and a change from TWSTRS Total of 53 at baseline to 38 for Week 4 (15 points improvement) for the subject in the 5000 U group, as this subject did have a Week 2 visit. In Study 302 the one subject was in the placebo group, and had no change from baseline imputed as the Week 4 outcome.

Because these were so few in number and not markedly imbalanced by treatment group, these missing data are not likely to have an important impact upon the interpretation of the studies.

TWSTRS Subscale Outcomes

The three subscales of the TWSTRS were examined individually. As for the TWSTRS Total, the mean baseline scores indicated considerably greater symptoms than the minimum required for eligibility. Baseline Severity was more than half of the maximal score, baseline Disability near, but less than half, and baseline Pain Subscale very near half the maximal.

Table 6 TWSTRS Severity Subscale Scores					
Time Point	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Baseline	18.4	20.2	20.2	22.1	22.6
Week 2	16.9	17.5	15.6	20.8	18.5
Week 4	16.2	17.0	15.4	21.0	18.9
Week 8	17.2	17.1	16.2	21.3	19.6
Week 12	17.4	18.8	17.9	21.5	20.7
Week 16	18.4	20.1	20.3	21.8	22.3

Table 7 TWSTRS Disability Subscale Scores					
Time Point	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Baseline	14.3	14.4	14.4	16.9	18.3
Week 2	13.2	11.9	12.5	16.3	14.6
Week 4	12.7	11.9	11.6	16.1	14.5
Week 8	13.1	12.7	12.2	16.4	15.1
Week 12	14.0	13.8	13.4	16.6	15.7
Week 16	13.9	13.7	15.1	16.5	16.4

Table 8 TWSTRS Pain Subscale Scores					
Time Point	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Baseline	10.9	11.8	12.4	12.2	11.9
Week 2	10.1	7.9	8.5	11.9	9.4
Week 4	10.4	8.2	8.2	12.1	8.4
Week 8	10.9	9.6	10.1	11.9	9.4
Week 12	10.8	1.04	11.5	12.4	10.5
Week 16	11.0	11.8	12.8	12.7	11.0

As more clearly summarized in the following table showing mean change from baseline, all three of the subscales contributed to the overall treatment effect. However, the most prominent component was the Pain Subscale, particularly when the difference from the placebo change is taken into account. Nonetheless, Severity and Disability subscales also were important to the overall effect.

Table 9 Improvement from Baseline in TWSTRS Total and Subscale Scores

TWSTRS Component	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Total	4.3	9.3	11.7	2.0	11.1
Severity Subscale	2.2	3.2	4.8	1.2	3.7
Disability Subscale	1.6	2.5	2.7	0.8	3.8
Pain Subscale	0.5	3.6	4.2	0.1	3.6

Uniformity Across Centers

Comment:

These were multicenter studies, and similar performance at multiple centers is an important characteristic of study results to support generalizability. These studies showed good uniformity of effectiveness of the toxin across the study sites in each study. Of the 9 sites in Study 301, only one site (# 8) indicated failure of the toxin to provide benefit in both dose groups, and this site was the source of the only result where a toxin group had an outcome inferior to the placebo group. Only at 2 other sites was one of the toxin groups not suggesting at least 2 points improvement of treatment effect, and these were distributed to one site's 5000 U group (site 23), and to another's 10000 U group (site 20). In Study 302's 7 sites, the toxin group was always numerically trending to improvement over placebo, and in only one site (# 19) was this by only 2 points. Overall, there was somewhat greater variability in the group outcomes in Study 302 than in Study 301. At two sites (# 9 and 18), toxin groups had notably large mean improvements while the placebo groups at those sites showed mean worsening of the symptoms; a phenomenon which was never seen in Study 301. However, taken as whole, the two studies suggest good reproducibility of the beneficial treatment effect across centers.

Table 10: TWSTRS Total Scores: Baseline and Improvement from Baseline at Week 4						
Site Code #	Time Point	Study 301			Study 302	
		Placebo	5000U	10000U	Placebo	10000U
All Sites	Baseline (n in group)	43.6 (36)	46.4 (36)	46.9 (37)	51.2 (38)	52.8 (39)
	Week 4 Improvement	4.3	9.3	11.7	2.0	11.1
4	Baseline (n in group)				45.0 (3)	51.8 (4)
	Week 4 Improvement				12.3	19.3
5	Baseline (n in group)	37.0 (4)	39.8 (4)	52.5 (4)	61.0 (5)	53.2 (5)
	Week 4 Improvement	2.8	11.5	15.0	5.4	8.8
8	Baseline (n in group)	45.2 (5)	55.0 (5)	56.3 (4)		
	Week 4 Improvement	4.2	0.6	4.8		
9	Baseline (n in group)	37.3 (4)	37.0 (4)	42.0 (4)	43.6 (5)	54.6 (5)
	Week 4 Improvement	3.5	5.5	13.0	-1.8	14.4
12	Baseline (n in group)	47.0 (3)	48.0 (2)	49.3 (3)		
	Week 4 Improvement	3.7	12.0	18.0		
17	Baseline (n in group)	52.0 (2)	49.0 (2)	47.0 (2)	58.0 (9)	53.3 (8)
	Week 4 Improvement	2.5	13.5	6.0	3.0	12.3
18	Baseline (n in group)				52.4 (5)	55.8 (5)
	Week 4 Improvement				-1.6	16.8
19	Baseline (n in group)	45.2 (5)	53.8 (5)	49.2 (5)	47.0 (6)	50.0 (6)
	Week 4 Improvement	3.0	6.6	13.2	3.8	5.5
20	Baseline (n in group)	46.0 (4)	45.8 (4)	41.3 (4)		
	Week 4 Improvement	5.8	15.8	6.3		
22	Baseline (n in group)	37.8 (6)	39.8 (6)	38.3 (7)		
	Week 4 Improvement	5.8	14.3	9.9		
23	Baseline (n in group)	54.7 (3)	51.0 (4)	53.0 (4)		
	Week 4 Improvement	6.7	7.5	19.0		
24	Baseline (n in group)				44.6 (5)	51.7 (6)
	Week 4 Improvement				-4.0	4.0

Empty cells indicate site did not participate in corresponding study

EFFICACY RESULTS: SECONDARY ENDPOINT OF PATIENT GLOBAL ASSESSMENT OF CHANGE

The prospectively stated secondary endpoint was the Patient Global Assessment of Change. This endpoint also supported the efficacy of the toxin treatment.

Comment:

Again there appears to be little additional benefit to the 10000 U dose over the 5000 U dose from this analysis. This includes examining the score out to later time points than 4 weeks, as well. At weeks 8 and 12, where there are higher scores in the toxin groups over the placebo group, the mean Patient Global score is nearly identical in the two toxin groups.

Table 11 Patient Global Assessment of Change VAS

Time Point	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
Week 2	47.2	59.1	60.4	38.3	56.0
Week 4	43.6	60.6 (0.001)	64.6 (0.0001)	39.5	60.2 (0.0001)
Week 8	46.9	50.6	51.0	40.9	56.2
Week 12	36.1	43.6	43.8	37.6	48.4
Week 16	37.3	35.9	34.5	36.2	43.7

p-values in () per ANOVA with investigator included, for pairwise comparison to placebo

Comment:

Similar to the case with the TWSTRS scores, the histogram of week 4 scores supports the conclusion of a general shift in patient outcomes to better scores. However, as with the TWSTRS changes, the histograms are difficult to interpret for amount of benefits, and cumulative plots of percentages of subjects with scores of at least specific global assessment amounts more clearly illustrates the breadth of the subject population's shift to better scores in the treatment groups. The curves for Study 301 also again clearly illustrate that the majority of the benefit was achieved with a dose of 5000 U. There is no prominent suggestion of improved benefit with 10000 U over 5000 U.

FIGURE 6

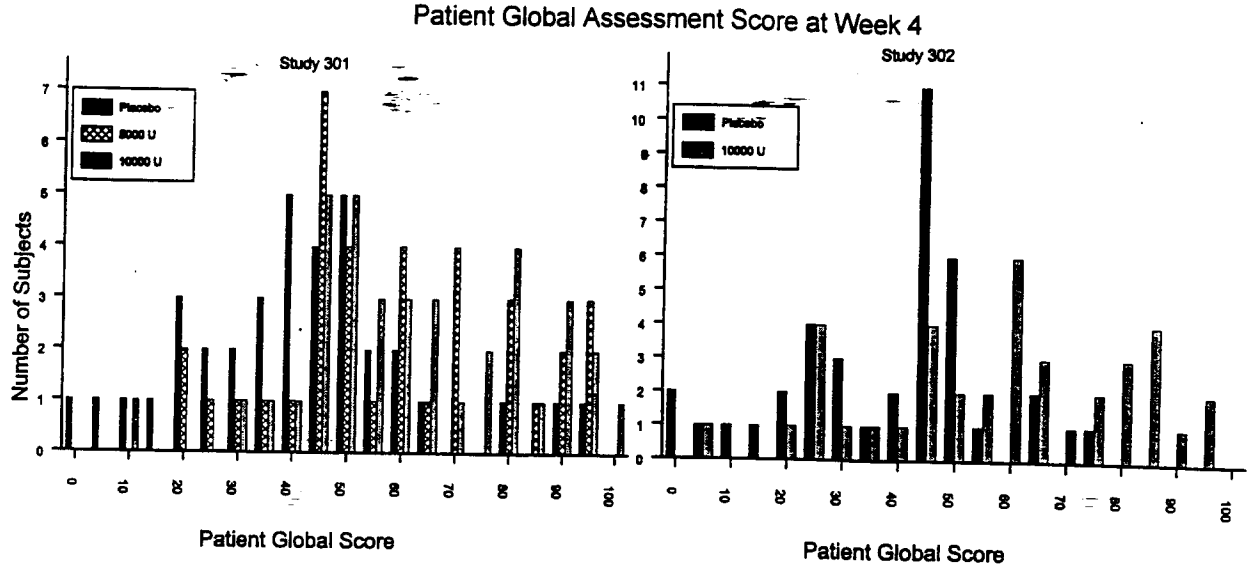
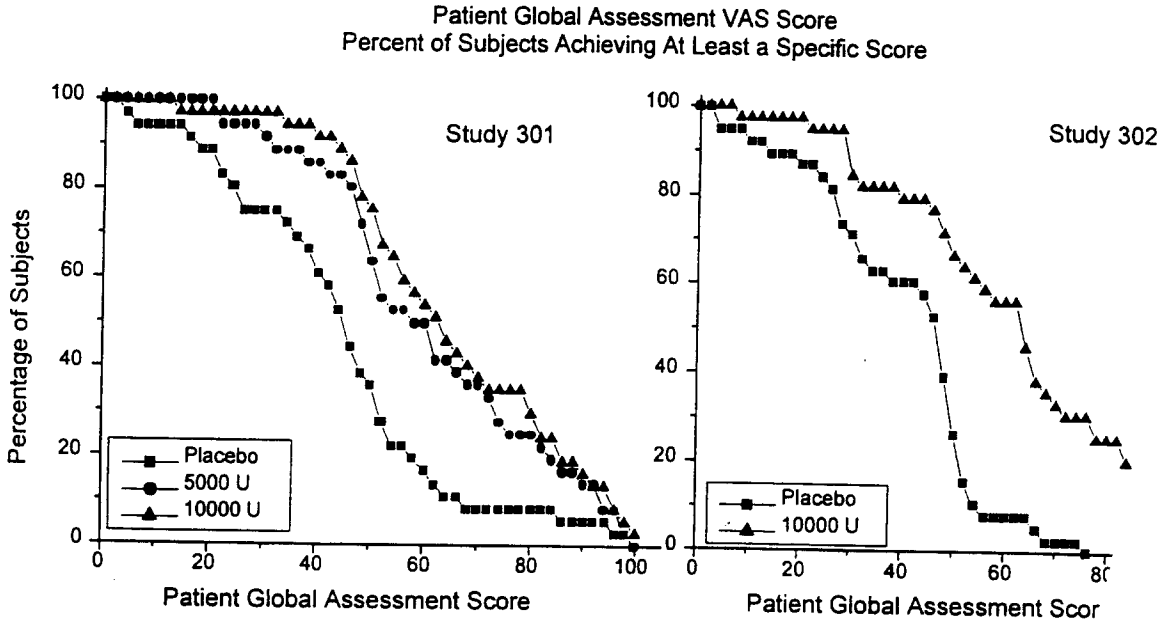


FIGURE 7



Unlike the TWSTRS, the Patient Global Assessment intrinsically incorporates a "ch baseline" nature to the assessment, as the question to which the subject responds ex directs to form a rating based on the subject's experience presently compared to ba

Comment:

Subjects in these studies were asked to complete the Patient Global Assess baseline visit (Day 0). The question asked at this visit was to score the sub compared to 1 week previous. This was different from the question attacl treatment visit VAS scales, where subjects were asked to score status con the Day 0 status.

Table 12	Patient Global Assessment of Change VAS - Baseline Response				
Parameter	Study 301			Study 302	
	Placebo n = 36	5000U n = 36	10000U n = 37	Placebo n = 38	10000U n = 39
# Obtained	36	36	37	36	37
Mean	39.0	35.3	42.4	34.5	42.5
Std. Dev.	21.5	20.7	18.0	19.9	18.1
Std. Err of Mn	3.6	3.4	3.0	3.3	3.0

This table is notable for the fact that in all groups, the mean response was of status being worse than 1 week previously. It is unlikely that during the 1 week period there was actually any significant change in the severity. Subjects were already months in time since the last botulinum toxin treatment, and the disease is generally stable over modest periods of time in subjects with these eligibility criteria. Thus, this suggests subjects perceived a decline in status that did not, on the whole, occur, and that subjects in the placebo groups were more inclined to perceive such a decline than in the 10000 U groups, particularly in Study 302.

EFFICACY RESULTS: OTHER EFFICACY ENDPOINTS

Two other important efficacy evaluations were conducted in the study, and were designated as tertiary endpoints. These are the Global Assessment performed by the investigator and the Patient Pain VAS assessment. These evaluations do not provide any uniquely new form of assessment. The Investigator Global is of course dependent upon the investigator's assessments that were obtained as part of the TWSTRS and comments from the patient that were likely influencing the Patient Global Assessment. The TWSTRS includes a pain subscale, and is thus likely to include many of the factors that influence the Patient Pain VAS, although not integrated into a score in the same manner. Thus they should complement the prior evaluations, and are expected to support their observed results, but not provide any new insights.

The Investigator Global Assessment of Change did fully support the observed effects on the primary and secondary endpoints. Botulinum Toxin Type B appears to provide efficacy in both studies at week 4. An additional observation is that the Investigator Global Assessment also provides no support for any increased amount of benefit with the dose of 10000U over the dose of 5000 U. The amount of benefit from toxin treatments, as assessed by the Investigator, is again indicated as relatively modest. A score of 50 is the midpoint of the scale, and is the implied value for no change from baseline. Thus, the best observed mean outcome, of 65, is only 15 mm on a 50mm scale segment ranging from no change to fully improved; the "average" subject received only about a 30% alleviation of their symptoms.

Table 13 Investigator Global Assessment of Change VAS

Time Point	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Week 2	54.5	62.5	63.8	49.7	61.0
Week 4	52.0	65.3 (0.001)	64.2 (0.004)	47.9	60.6 (0.0001)
Week 8	51.1	59.2	59.7	48.8	58.9
Week 12	48.7	49.3	54.1	45.2	51.8
Week 16	47.0	48.7	45.7	46.8	49.3

p-values in ( ), per ANOVA with site included, for pairwise comparison to placebo

The Patient Pain VAS scores, shown below in Table XX, also support the previous observations. Both studies indicate efficacy was associated with toxin treatment. The Patient Pain VAS in Study 301 again indicates that there was little, if any, increased benefit to the dose of 10000 U over the dose of 5000 U. Both studies again indicate that the amount of benefit is modest in absolute terms. Most subjects, at peak effect, remained with substantial amounts of pain.

Comment:

However, this assessment is difficult to quantitatively assess. The placebo and 10000 U groups were not equivalent in mean pain score at baseline. The 10000 U group assessed their pain as worse than the placebo group did in Study 301, and the reverse was true in the Study 302. The placebo groups in both studies had mean pain scores that were generally constant, while all three toxin groups showed improvement post treatment, and then return of pain in later weeks. In Study 301, where the amount of improvement from baseline was larger than in Study 302, the toxin treated groups reported worse pain at week 16 than they had at baseline. The importance of the baseline differences upon the subsequent responses is not known (e.g., subjects who report more pain may have a different likelihood to report greater improvement due to placebo effect; subjects who have more pain may be more readily relieved of pain by treatment than those with lesser amounts of pain, etc.). All of these factors leave this evaluation as one that is fully consistent with the prior assessments, but adding little additional strength to the evidence. Designation of this endpoint as a tertiary endpoint was an appropriate choice by the study sponsor.

Table 14 Patient Pain VAS					
Time Point	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Baseline	43.6	40.8	35.1	33.6	41.4
Week 2	43.7	63.7	59.8	38.1	52.5
Week 4	43.7	61.7	62.3	37.3	57.7
Week 4 - Change from Baseline	0.1	20.9 (0.001)	27.2 (0.0002)	3.7	16.3 (0.001)
Week 8	46.8	53.2	49.3	38.4	52.1
Week 12	39.8	44.3	40.2	33.7	43.6
Week 16	41.4	33.0	28.4	31.0	42.2

p-values in ( ), per ANCOVA with investigator and baseline score included, pairwise comparison to placebo

An additional note on this assessment tool is that unlike the TWSTRS, there was no minimum score in Patient Pain VAS as an entry criterion. Thus, subjects with baseline Pain VAS of nearly no pain were eligible, and these subjects had little or no opportunity to benefit on study in this assessment. The number of subjects with baseline Pain VAS greater than 90 was 0 or 1 in each treatment group except for the 5000 U group in Study 301, which had 3. The number of subjects with baseline score greater than 95 was 0 in both placebo groups, 1 in both 10000 U groups, and 2 in the 5000 U group. Thus, there were too few patients for whom this ceiling effect might exist, and they were largely well balanced among the groups, for this to have any important effect on the study outcome.

Analysis of Subjects With At Least 20% Improvement

An additional analysis approach proposed by Elan, and eventually relegated to a tertiary endpoint after extensive discussions with CBER was a responder analysis. Elan defined responders as subjects who showed a 20% improvement from baseline on the Week 4 TWSTRS Total score. Their analysis on this basis is seen in Table XX.

Table 15 Subjects with TWSTRS Improvement by at least 20% of Baseline					
Parameter	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Number with 20% Improvement	10	18	22	3	17
% with Improvement	28	50	59	8	44
p-value		0.051	0.01		0.001

p-values for pairwise comparison to placebo group

The basis of the selection of the 20% criterion is uncertain, but was prospective. After CBER repeatedly expressed concern with the clinical meaningfulness of this analysis, Elan assembled a panel of applicant selected investigators to discuss the criterion in December 1997. This group expressed an opinion that defining a clinically meaningful response to a treatment in general should be determined by both the patient self-assessment and the physician clinical evaluation. The panel supported the use of the TWSTRS as the primary outcome of a study. The panel agreed that on the Global VAS, any score of greater than 50 indicates an indication of improvement [Comment: However, no data supporting the reliability of this scale, which has no visual marker of the no change point was provided or considered. Thus, this assumption is unsupported.] The panel members agreed that there were strong correlations between TWSTRS Total and Patient Global VAS, and what they subjectively consider a clinically meaningful effect, based on their experience. [Comment: Again, no quantitative data were examined to support the assumption of strong correlations. Therefore, this opinion remains unsupported.] The panel noted that TWSTRS was not a uniform interval scale; 5 or 15 points change at the low end of the scale may not mean the same to a patient as 5 or 15 points change from a higher baseline TWSTRS. The panel proposed that relative change is more meaningful than absolute change. The panel reviewed the recommendations of another (uncertain composition) panel that met in July 1993 and recommended a 20% change as meaningful, based upon experience (no data was referenced). This panel endorsed that criterion, based in part upon that criterion giving the expected response rates of 65% treated, 30% placebo in Study 009 [Comment: The basis for the expected response rates for an effective therapy are unstated, and seem to be irrelevant to the issue of a meaningful change.] , and that 20% appeared to be an appropriate change to be meaningful, in their opinion and experience [Comment: Again, no data was examined or provided to arrive at this opinion.].

However, a final recommendation of the panel was that a patient who had both a 20% and a positive self-assessment would be recommended for retreatment [Comment: i.e., 20% change alone would not lead to retreatment; patient self-assessment showing an effect would also be required. Thus, the panel implicitly expressed doubt as to the meaningfulness of a 20% change in the absence of any other supportive data.].

In summary, while a panel of investigators was selected by the applicant, they were unable to provide any clear, data based, rationale to support the 20% criterion as meaningful. Because of this weakness in the interpretation, expressed by CBER, the endpoint was relegated to a tertiary endpoint by Elan. Further examination of this analytic approach is deferred to a following section.

EFFICACY RESULTS: SUMMARY OF MAIN EFFICACY RESULTS

Table 16 summarizes the most important efficacy results from these two studies, shown as estimated treatment effect size, confidence interval on that treatment effect, and the p-value for the treatment effect, for each toxin dose. All scoring systems have been calculated so that positive effect size indicates improvement (i.e., TWSTRS scores calculated as [baseline - outcome] due to the fact that improvement on TWSTRS entails a lowering of the score). Comparisons of toxin to placebo treatment effect at week 4 for all of the assessment tools were statistically significant for all three of the toxin treatment groups.

Table 16 Comprehensive Summary of Efficacy Results

Endpoint	Study 301				Study 302	
	5000 U n = 36		10000 U n = 37		10000 U n = 39	
	Est Tx Effect	95% CI p-value	Est Tx Effect	95% CI p-value	Est Tx Effect	95% CI p-value
TWSTRS Total Week 4	5.0	8.9, 1.2 0.012	7.2	11.1, 3.3 0.0004	8.7	12.2, 5.2 0.0001
Patient Global Assessment Week 4	17.0	7.0, 26.9 0.001	21.2	11.3, 31.1 0.0001	20.1	11.2, 29.1 0.0001
Investigator Global Assessment - Week 4	13.4	5.5, 21.3 0.001	11.8	3.9, 19.7 0.004	12.7	7.4, 18.1 0.0001
Patient Pain VAS - Wk 4	19.2	8.0, 30.4 0.001	21.8	10.5, 33.0 0.0002	15.9	6.7, 25.2 0.0001
TWSTRS Severity Subscale - Wk 4	0.9	2.5, -0.6 0.22	2.5	4.0, 1.0 0.0016	2.4	3.9, 1.0 0.001
TWSTRS Disability Subscale - Wk 4	1.0	2.7, -0.7 0.26	1.1	2.8, -0.6 0.19	2.5	4.1, 1.0 0.002
TWSTRS Pain Subscale Wk 4	2.9	4.7, 1.1 0.002	3.2	5.1, 1.4 0.0008	3.5	5.0, 2.1 0.0001
TWSTRS Total Week 8	4.8	8.8, 0.8 0.019	6.1	10.1, 2.1 0.003	6.8	10.2, 3.4 0.0002
TWSTRS Total Week 12	2.0	6.0, -2.1 0.33	2.5	6.5, -1.6 0.22	4.8	8.5, 1.0 0.013

p-values per ANCOVA for TWSTRS scores and Pain VAS, ANOVA for Global Assessments
p-values for pairwise comparisons to placebo

Comment:

Within the TWSTRS scores, this table suggests that the toxin's benefits are substantially weighted to reduction in the pain of cervical dystonia. Of the total effect size on the TWSTRS Total, the Pain Subscale is always the largest contributor of the three subscales. Only the Pain Subscale showed a statistically significant treatment effect consistently. While this does not demonstrate that there was no meaningful effect on the other aspects of the disease, it does suggest where the most prominent effect is.

EFFICACY RESULTS: EXPLORATORY ANALYSES OF EFFICACY ENDPOINTS

Correlation of Change in TWSTRS with Global Assessments

Comment:

In order to further address the divergence of viewpoints regarding meaningfulness of any specific criterion of TWSTRS change, correlations between the TWSTRS Scale and the other outcome scales were examined to better understand the behavior of these outcome assessments, and potentially the meaningfulness of specific amounts of change.

However, there was only limited correlation between the change in TWSTRS Score to Week 4 and the Global Assessments at Week 4 or the change in the Pain VAS at week 4. This suggests that the CD symptoms are being integrated into the patient's awareness and self-assessment of their condition in a manner that is notably different than the formal system devised in TWSTRS.

Table 17: Correlation Coefficients of Outcome Scales at Week 4

	Patient Global Assessment	Investigator Global Assessment	Change in Pain VAS
Change in TWSTRS Total Score	0.60	0.69	0.57
Percent change in TWSTRS Total Score	0.58	0.73	

n = 186 observations for each correlation calculation

This low correlation suggests that modest changes in TWSTRS would have very variable Patient Global Assessments associated with them (as well as Investigator Global or Patient Pain VAS). Scatterplots of these scores against each other confirm this.

An approach to an improved responder analysis is to use a dual criterion; only patients with a stated amount of TWSTRS improvement and self-recognition of the improvement (as indicated on the Patient Global VAS) are scored as a responder. One such analysis is to designate as a responder only subjects with both TWSTRS improvement of 20% or more and Patient Global VAS score of 65 or more. This is shown in the following table.

Table 18 Post hoc Responder Analysis:
Subjects with TWSTRS Improvement by at least 20% of Baseline And Patient Global VAS > 65

Parameter	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Number with 20% Improvement	2	11	13	1	11
% with Improvement	5	31	35	3	28
p-value		0.012	0.003		0.003

p-values for pairwise comparison to placebo group, Fisher Exact test

This post hoc analysis clearly suggests that there are more responders in all three toxin treated groups than their respective comparators. Relaxing the criterion to those subjects with 20% TWSTRS improvement and Patient Global of >60 only increases the number of responders slightly in the toxin groups, suggesting that the specific value selected for the second criterion is less important than using it to eliminate subjects who have TWSTRS changes, but do not themselves recognize any improvement.

A responder analysis is fully consistent with the primary and secondary endpoints, but the post hoc nature of these analyses and lack of sufficient rationale for the prospective plan makes any responder analysis untenable for a definitive description of the fraction of subjects who benefit from the treatment.

Correlation of TWSTRS Pain Subscale with Patient Pain VAS Assessment

The limited correlation of change in TWSTRS Score with change in the Pain VAS was not substantially improved when the either the Pain Subscale in TWSTRS was examined (r = 0.64) or the Pain Severity subsection of the TWSTRS Pain Subscale was examined (r = 0.64) and compared with the Patient Pain VAS. This implies that the explicitly defined manner in which the TWSTRS Pain Subscale is examining subject pain is different from the manner that subjects integrate their pain experience when responding to the Pain VAS question. There was no inconsistency between the two approaches with regards to toxin treatment effect; both approaches indicated that benefit is derived from treatment with toxin. However, the lack of correlation suggests that the formalized approach used by TWSTRS does not weight factors, incorporate all factors considered by the patient, or both.

Pain Medication Usage

During the studies subjects were to continue analgesic medications required on a regular basis at baseline, and were discouraged from initiating new analgesic medications. If new analgesics were initiated, however, subjects were to discontinue the analgesic 1 day prior to the scheduled evaluations so as to not interfere with the pain components of the evaluations.

Responding to a requested analysis from CBER, Elan tabulated the subjects who used analgesic medications during the study. The majority of subjects used some pain medications at some time during the study. Elan analyzed the pain medication use in two manners. One was by use of any medication that could provide analgesia, and the second was for "strict" analgesic use, defined as those medications used for stated purpose of analgesia and which might not have entirely washed out in the 24 hr abstinence period prior to the scheduled evaluations (i.e., those medications that might have biased the outcome assessments).

Table 19: Number of Subjects Using Analgesics during Study					
	Study 301			Study 302	
	Placebo n = 36	5000 U n = 36	10000 U n = 37	Placebo n = 38	10000 U n = 39
Any Analgesic	27	26	28	29	31
Narcotic Type Analgesic	8	7	8	8	7
Strict Analgesic by Week 4	18	11	11	12	16
Strict Narcotic by Week 4	5	2	2	5	3

These analyses indicate that there was not an imbalance in total use of analgesic type medications. The categories of strict definition show some modest differences in relative subject numbers, but these are still a minority of subjects and most imbalances are to larger numbers of placebo subjects using medications. If these had biased the studies, they would lead to a falsely small treatment effect estimate. Elan concludes that analgesic use could not have significantly biased the week 4 outcomes of the study as there was sufficient balance of analgesic use across treatment groups.

Duration of Effect Analyses

The issue of estimating the duration of a treatment effect was raised by CBER with Elan during design of the studies. Elan responded that the primary outcome of the studies should be used to assess the duration of the effect, and should be through an analysis that includes all subjects, not a subset of subjects based on treatment effect. Elan elected to conduct a Kaplan-Meier survival curve analysis on subjects for time to return to baseline TWSTRS Score. A recognized difficulty with this approach is that the post treatment assessments were spaced at wide intervals relative to the expected duration of effect. This will markedly decrease the ability to precisely characterize the duration of effect by a time to return to baseline method. However, in favor of this approach, Elan noted that clinical practice is to usually reinject the muscles when the patient returns to their baseline condition.

For Study 301, Elan found that median time for return to baseline was 9, 16, 16 weeks respectively in the placebo, 5000 U, and 10000 U groups. In study 302 the median times were 8 and 16 weeks for the placebo and 10000 U groups. Since the assessments were at weeks 12 and 16, this suggests that the actual median time to return to baseline for BotTx-B treated subjects was between 12 and 16 weeks. This analysis also fails to provide any notable evidence of important differences between the 5000U and 10000 U doses.

A second approach to this question was based upon a responder subset analysis. In Study 301 there were 18 subjects in the 5000 U group and 22 in the 10000 U group who showed a response of at least 20% of baseline; these were included in this analysis. Grouped together, most (34, or

85%) of these 40 subjects had not returned to baseline by week 12, while most (29 or 72%) had returned to baseline by 15 weeks. In Study 302 there were 17 subjects in the 10000 U group who were classified as responders at week 4. Only 1 subject had returned to baseline by week 12 (94% not returning to baseline), while only 6 subjects (35%) had not returned to baseline by week 16. These alternative analyses are consistent with Elan's survival analyses of the overall study population suggesting that responses will have a median duration between 12 and 16 weeks.

Antibody Formation Results

The potential for antibody formation to occur is clearly present in this therapy. There is both risk of adverse events upon readministration in presence of antibody as well as potential for neutralization of the toxin prior to activity upon the nerve terminals, and thus loss of efficacy.

The subjects in these single dose studies had not previously received BotTx-B so that neither was a reasonable possibility in these studies. However, the rate of antibody formation observed in studies remains an indication of the potential for these effects to occur with repetitive use.

Subjects had antibody testing performed on baseline, week 4 and termination (week 16 in most subjects) serum samples. Testing was performed with a screening ELISA assay, and a mouse neutralization assay on subjects who tested positive in the ELISA.

For Study 301 there were 108 subjects with samples monitored. There were only 15 subjects with a positive ELISA on at least one sample

Table 20: Numbers of Subjects with Positive Antibody ELISA Results

	Study 301		Study 302	
	Placebo	5000 U or 10000 U	Placebo	10000 U
Baseline	3	8	3	5
Week 4	4	10	4	5
End of Study	4	9	4	4

Note: Total number of samples assayed not clearly stated in submission
Study 301 results given solely as combined toxin groups

Most subjects with positive titers had titers of 20 units/ml or less. The notable exception was one toxin patient in study 301 who entered the study with a titer of 75, and remained with titers of 60 and 66 at the Week 4 and Termination follow-ups.

Comment:

However, not all subjects had samples tested at all time points. Elan did not provide an analysis of numbers of samples tested for the positive ELISA it obtained.

Of these ELISA positive samples where enough material was available to be tested in the mouse neutralization assay, none were positive for neutralizing antibodies. This was not surprising, as

the ELISA titers were generally low, and likely to be not detectable by the mouse assay, even if they were neutralizing antibodies. Thus, none of these subjects would be expected to be at risk of immune mediated adverse events or loss of efficacy upon the next repeat injection.

However, these results were obtained only after a single, first dosing session. These studies may not be predictive of antibody formation with long term, repetitive use.

SAFETY RESULTS

Deaths and Serious Adverse Events

There was solely one death between the two studies, which was also the only withdrawal due to adverse event. This was a patient with preexisting known coronary artery disease who suffered a myocardial infarct, and subsequently died from this event.

There were 6 additional subjects in Study 301 who had a serious adverse event (SAE). In the placebo group, 2; with events of bladder cancer and atrial flutter. Two in the 5000 U group, cases of coronary occlusion and bladder stenosis; and two in the 10000 U group with diverticulitis and a fracture of the foot due to a motor vehicle accident.

There were a total of 5 SAE in Study 302. These included cases (1 each) of basal cell skin cancer, rectal cancer, gastric ulcer (recurrent), periappendiceal abscess, and angina. None of the SAE, in either study, were deemed related to study agent.

There was one withdrawal due to adverse event in Study 302. A 75 yo woman had worsening cervical dystonia symptoms 4 days after study treatment, and chose to withdraw from further study follow-up. This worsening was not classified as a SAE.

Adverse Events of All Natures

A summary tabulation of adverse events of all types in subjects in the study is shown below in Table 21. Only event types where the incidence was at least 10% in one of the toxin treatment groups within each study is shown. Although 10% incidence is a substantial rate, these studies were small and lower rates constitute just 3 subjects.

The most prominent events associated with toxin injections were dry mouth and dysphagia. There is suggested also a modest increase in injection site pain and dyspepsia. The other events included in the table do not appear to have a notable or consistent increased incidence in the toxin groups.

Table 21 Percentage Incidence of Most Frequent Adverse Events

Adverse Event	Study 301			Study 302	
	Placebo n = 36	5000 n= 36	10000U n = 37	Placebo n = 38	10000U n = 39
Injection Site Pain	8	6	11	8	18
Infection (unspecified)	28	25	8	16	21
Headache	8	25	14		
Pain	14	6	24		
Arthralgia				3	13
Pain related to CD	25	31	27	21	21
Dry Mouth	3	14	24	3	44
Dyspepsia	8	3	11	5	13
Dysphagia	3	11	19	5	28
Nausea				8	15

Note that each 3% is just 1 subject in any group; 10% is 4 subjects
Blank cells indicate that fewer than 10% in any treatment arm in the study had that event

Due to the relatively small number of subjects with most AES, discussion of the distribution of intensity is deferred to later in the review.

Dysphagia

Dysphagia is the most common important adverse effect of BOTOX treatments for cervical dystonia reported in the medical literature. It is not surprising that it occurs with increased frequency associated with BotTx-B treatments.

In Study 301 there were 13 reports of dysphagia. None were severe or serious, 3 of moderate severity and 10 were mild. All but 1 resolved, and the non-resolving case of dysphagia was in a subject who entered the study with a report of dysphagia as a chronic condition which increased in severity on study. These reports include at least several subjects who reported a change in their diet content for several days due to the event. No subjects required hospitalization or tube feeding for the event.

In Study 302 there were a total of 13 dysphagia reports as well, 2 in the placebo group (1 mild, 1 moderate), and 11 in the toxin group (6 mild, 5 moderate). Thus the incidence was markedly increased in the toxin treated groups, but there were no cases of severe dysphagia reported in either of these two studies.

In response to CBER raising the issue during study design, Elan performed analyses examining for an association of adverse events with specific muscle dosing. For their preferred analysis, they grouped subjects by approximate doubling of dosing for each muscle (approximately

tertiles) and examined the AE incidence in the subjects without toxin injection to the specific muscle (placebo included) and each tertile group (i.e., none, low, medium, highest dose groups by specific muscle).

The trapezius muscle in Study 301 appeared to show a correlation for dry mouth (4%, 27%, 31%, 63%) and for dysphagia (5%, 18%, 23%, 38%) with dose-received groupings (mean dose in tertiles 1000 U, 2115 U, 4375 U), but the tertiles were small in size (n= 11, 13, 8).

The association in Study 301 was also suggested for the sternocleidomastoid muscle for dry mouth (6%, 15%, 19%, 31%) and for dysphagia (2%, 15%, 19%, 25%) for tertiles with mean doses of 850 U, 1560 U, 3344 U. The SCM is more frequently involved muscle in CD, and the tertiles were larger for this case, n = 20, 21, 16 .

The Study 301 levator scapulae muscle analysis also suggested an association of dosing with dry mouth of similar slope as the SCM, but there were smaller numbers of subjects. Other muscles did not provide suggestions of associations.

Study 302 had fewer toxin injected subjects, and thus the individual muscle tertiles were smaller and less reliable. In Study 302, the trapezius muscle was injected in only 10 subjects, and this analysis was not informative. The SCM muscle had tertiles of n = 11, 11, 11 subjects with mean doses of 1818 U, 2936 U, and 4305 U, and did suggest an association with dysphagia (7%, 18%, 36%, 36%). Elan reports that no other associations were noted.

However, these analyses entailed considerable subdivisions of subjects and combining into arbitrary groups. In response to a CBER request during the analytic plan prospective discussion, Elan also performed logistic regression analysis of the incidence of dysphagia with dose in selected muscles. These analyses confirmed a dose related increase in the incidence of dysphagia with SCM injections. Severity of dysphagia was not significantly dose related, but as few subjects developed dysphagia, there were wide confidence intervals on this analysis. The combined studies analysis suggests a point estimate of incidence of dysphagia with 5000 U of toxin injected into the SCM at approximately 60%. Other muscles were not examined with this method.

PHASE 2 STUDY AN072-009

OVERVIEW

Study AN072-009, hereafter referred to as Study 009, was a late stage phase 2 study that provided preliminary evidence of efficacy. Eligibility criteria and treatment plan were similar to that already described for the phase 3 studies. This study was most different in requiring follow-up evaluations only through week 4, after which only subjects who had treatment benefit were continued in follow-up assessments. This study is important in providing a broader range of doses studied within a single trial, with sample size per group almost as large as the phase 3 studies.

Title: A double-blind, placebo-controlled, single treatment, dose-finding, safety, tolerability and preliminary efficacy study of BotB (Botulinum toxin type B) at various doses in patients with idiopathic cervical dystonia.

The original version of the protocol was finalized January 1995. Amendments 1, 2,3 were made before study initiation and are also included in this description. Amendment 2 changed the dose levels used from the initially planned levels of 2400, 5000, and 7500 U to the actually used levels of 2500, 5000, and 10000 U.

CLINICAL STUDY DESIGN

Objectives

The stated objectives of this study were:

To evaluate the safety and tolerability of three doses of BotB in patients with CD
To provide dose-finding information to aid in selection of effective doses to be tested in future studies
To assess the performance of the TWSTRS, TWSTRS subscales, and a video taping evaluation process

General Design Structure

This was a multicenter, double blind, placebo controlled study of a single treatment with study agent in a 4 parallel group design of three doses and placebo. Subjects receive treatment into 2-4 affected cervical muscles at the one treatment session, and return for follow up evaluations at weeks 2 and 4. At week 4 the primary outcome evaluations are conducted. Those patients who appear to show improvement over the baseline evaluation will continue to have follow-up visits every 4 weeks. Subjects were terminated from the study at week 4 for failing to show the 20% improvement from baseline, or when their response had declined by half from the week 4 effect.

A total of 120 subjects were planned for this study, with even randomization between the 4 groups, and participation of 15 centers. All subjects were those that had previously been treated

with Botulinum toxin type A and been considered to have a worthwhile response. Subjects that had secondarily become inadequately responsive to type A toxin were permitted to enroll into the study, but were required to be less than 50% of the enrollment at each site. Randomization was not stratified by this factor.

Eligibility Criteria

Inclusion Criteria

1) Diagnosis of idiopathic cervical dystonia, of at least 1 year duration, but not more than 10, involving at least 2 of the 6 specified cervical muscles.
2) CD of at least moderate severity, defined on the TWSTRS scale at baseline by Total >= 20 ; Severity >= 10 ; Disability >= 3 ; Pain >= 1
3) Male or Female, >= 18 yo
4) Body weight >= 46 kg

Exclusion Criteria

1) Subject has never had a good response to prior botulinum toxin injections of any kind
2) Any botulinum toxin injection within prior 4 months
3) Not returned to baseline following their last treatment with botulinum toxin
4) Pure retrocollis or anterocollis, or contractures of muscles such that neck ROM decreased
5) Pregnancy, breast feeding, inadequate contraception (women)
6) Irregular use of drugs for symptomatic treatment of CD
7) Myotomy or denervation surgery for treatment of CD
8) Use of aminoglycoside within 30 days prior to screening
9) History of other significant neurological disorder that may predispose to increased risk
10) Other general medical disorders leading to increased risk or decreased ability to complete evaluations

Subjects were classified as resistant to Type-A toxin if subject has a prior history of good response to Type A toxin, failed to have good response on two successive treatment sessions with adequate toxin amount, one of which was an increase in dose. Each site was permitted to have not more than 50% of enrollment at any time consist of Type A resistant patients.

Study Treatments

Subjects were randomized to 4 treatment groups: Placebo, 2500 U, 5000 U, 10000 U.

Toxin was supplied in vials containing either 2500 or 5000 U of BotTx-B, in 1 ml of buffer of 0.5mg HSA, 2.7mg NaSuccinate, 5.8mg NaCl (this differed from the phase 3 study supplies in having a higher amount of NaSuccinate). Placebo vials contained 1ml of buffer only. The toxin used in this study was from Batch A1, which had been manufactured in the [redacted] facility.

Study treatment was shipped as boxes with two vials, to be kept refrigerated until use. Each vial box had the sequential patient identifier number on the box. Sites would use the boxes in

sequential order as subjects were enrolled. The contents of the two vials were combined to make 2ml of total study agent for each subject. By mixing placebo, 2500, 5000 U vials in the boxes appropriately, the 4 treatment group doses were provided in a blinded manner to the study site. Investigators could further dilute the study treatment to up to 5ml volume with 0.9% sterile saline if they preferred for the study injections.

Investigators selected 2 to 4 muscles for injection, and determined the dose per muscle as desired by determining the number of injection sites for each muscle, with the total dose required to be the total assigned volume. Each injection site could receive no more than 1ml nor less than 0.1 ml of volume.

Subject Evaluations

Medical history, physical exam, neurological specific history and exams were performed at study start and end of participation for each subject. Vital signs, clinical laboratory assessments, adverse event monitoring, and concomitant medication recording were performed throughout the study period at subject visits to the study site. A phone call to the subject 1 week after the study injections for adverse event monitoring was also performed. Serum samples for antibody testing were obtained at baseline and week 4 only; but not at the termination visit. In addition, study assessments relating to efficacy included the following:

Screening Visit
- TWSTRS Evaluation
- Patient Pain VAS

Day 0 (Baseline, with study treatment post evaluations)
- Neck muscle size assessment,
- TWSTRS evaluation, in person and on videotape
- Patient Pain VAS
- Baseline assessment for Investigator Global Assessment
- SIP

Week 2 and Week 4
- Neck muscle size assessment
- TWSTRS in person; on videotape at Week 4 also
- Patient Pain VAS
- Physician and Patient Global Assessment VAS
- SIP at week 4

Every 4 weeks until termination (If Subject had >= 20% of baseline TWSTRS improvement)
(Weeks 8, 12, 16, as needed)
- TWSTRS
- Patient Pain VAS
- Patient and Investigator Global Assessment VAS

Visit Scheduling: The week 1 phone call, and the Weeks 2 and 4 visits were required to be within 3 days of scheduled date. Further follow-up visits were to be within 7 days of the scheduled date.

Evaluation tools were as previously described with the phase 3 studies.

Endpoints and Planned Analyses

Efficacy Endpoints

Primary: TWSTRS Total at Week 4.

Secondary:

TWSTRS Subscales: Severity, Disability, and Pain Subscale
Patient Pain VAS
Patient Global VAS
Investigator Global Assessment VAS
Percentage of Responders

Analytic Plan

A separate Analytic Plan document was written, but not dated. While more detailed, it was not substantially different than the analytic plan described within the protocol.

Two datasets were used for analysis. A dataset called "Intent to Treat" consisted of all subjects who received study drug and had at least 1 visit with efficacy data recorded. Several evaluable subject datasets were also described.

The Primary Endpoint was evaluated with an overall test for group differences, tested with ANCOVA including independent variables of center, treatment group and baseline TWSTRS. The primary analytic dataset was stated as subjects who have baseline and Week 4 assessments. No provisions were made for imputation of missing data.

Secondary endpoints also analyzed with a similar ANCOVA method. Pairwise comparisons of each dose group with placebo was also planned to be done.

An Interim Administrative Analysis was planned when 50% of subjects had completed the week 4 visit. The purpose was to facilitate planning of future studies, and no change planned in the conduct of this study. No statistical adjustment was planned.

Median time to loss of response will be estimated for each group with Kaplan-Meier analysis.

TWSTRS scores were to be derived from videotape assessments as well, and compared for agreement with the in-person scores.

Planned Sample Size

The sample size required was estimated as 30 subjects per each of 4 groups to provide 90% power to detect treatment effect of 7 points in the high dose group when sd is 7 points, and two sided test of overall treatment effect is applied at 0.05 level.

STUDY PERFORMANCE AND SUBJECT DISPOSITION

Enrollment and Subject Disposition

There were 122 subjects enrolled into the study, which initiated in September 1995 and completed April 1996. Only 3 subjects were deemed non-evaluable for the Evaluable subject datasets (1 each due to duration of CD < 1yr, no TWSTRS videotape at week 4, and incomplete videotapes), so that there was no substantial difference between the evaluable and ITT datasets.

No subjects were in fact excluded from the "modified Intent to Treat" analysis set for the Week 4 evaluation, so that it can be treated as a true intent to treat dataset.

Mean time on study was shorter in the placebo group than in any of the other groups, 45 days in placebo group, 61, 67 and 75 days respectively in the 2500, 5000, and 10000 U groups. This difference is reflective of the difference between the groups in fraction of subjects with an observed response.

There were 12 study sites participating in the study. The distribution of subjects between groups within each site was very even, likely indicating a small blocking size for the study within each center. Most sites had 2 or 3 subjects in each group.

Treatment Administered

Comment:

Elan has not submitted information characterizing the details of the dosing by muscle.

Demographics and Baseline Characteristics

The study population characteristics were in general similar to those patients that had been enrolled into the phase 3 studies, and already reviewed. There was good comparability between the treatment groups for most parameters.

Table 22: Demographic and Baseline Characteristics - Study 009				
Parameter	Placebo n = 30	2500 U n = 31	5000U n = 31	10000U n = 30
Age (mean, yrs)	57	54.1	52.4	56.5
Sex % female % male	70 30	68 32	71 29	60 40
Race % White %black % hispanic % other	100 0 0 3	100 0 0 0	97 0 3 0	90 7 3 0
Height (cm, mean)	168	167	168	168
Wt (kg, mean)	74	75	75	71
Baseline TWSTRS Total	45.5	45.6	45.2	47.5
Baseline Patient Pain VAS	52.0	45.8	49.7	46.2

EFFICACY RESULTS: PRIMARY ENDPOINT AND SUBSCALES

TWSTRS Scores and Changes from Baseline

The overall test of the primary endpoint showed a statistically significant difference with p=0.0001.

Specific results at the primary endpoint time of week 4, as well as other visits are shown in the following tables. Note that non-responder subjects were intended to not be evaluated beyond week 4, so that there are substantially fewer numbers of subjects with follow-up evaluations at successively later timepoints. Thus, any post Week 4 data becomes difficult to interpret because of the markedly fewer subjects.

Table 23 TWSTRS Total Mean Scores - Study 009				
Time Point	Placebo n = 30 enrl	2500 n= 31 enrl	5000U n = 31 enrl	10000U n = 30 enrl
Baseline	(30) 45.5	(31) 45.6	(31) 45.2	(30) 47.5
Week 2	(30) 42.5	(30) 34.9	(30) 35.6	(30) 32.1
Week 4	(30) 42.2	(31) 34.0	(31) 32.7	(30) 31.1
Week 8	(8) 36.4	(18) 33.9	(18) 32.7	(23) 31.9
Week 12	(5) 26.6	(11) 39.2	(14) 35.3	(16) 38.1
Week 16	(3) 26.3	(5) 42.4	(9) 32.8	(9) 41.9

Number enrolled given at top of column
Number evaluated at timepoint in ( )

When examined as a change from baseline in the TWSTRS Total score, the changes at week 4 for all three toxin groups was substantially more than for placebo, and were all statistically significant in pairwise comparisons with placebo.

Table 24 TWSTRS Total Scores Change from Baseline - Study 009				
Time Point	Placebo n = 30 enrl	2500 U n= 31 enrl	5000 U n = 31 enrl	10000 U n = 30 enrl
Week 2	(30) 3.0	(30) 9.8	(30) 9.2	(30) 15.4
Week 4 p-value	(30) 3.3	(31) 11.6 (0.002)	(31) 12.5 (0.0005)	(30) 16.4 (0.0001)
Week 8	(8) 6.1	(18) 12.9	(18) 13.4	(23) 15.3
Week 12	(5) 12.8	(11) 11.4	(14) 12.5	(16) 10.8
Week 16	(3) 15.0	(5) 6.2	(9) 12.2	(9) 8.6

Number evaluated at timepoint given in ( )
P-value given in ( ) in second line for Week 4 outcome only

The comparison between the treatment groups is also clearly shown in the curves of cumulative percent of subjects achieving specific amounts of change (in either TWSTRS points or percent of baseline score). These figures suggest that the majority of the treatment effect is available at the lower doses, and the highest dose adds little additional treatment effect.

FIGURE 8

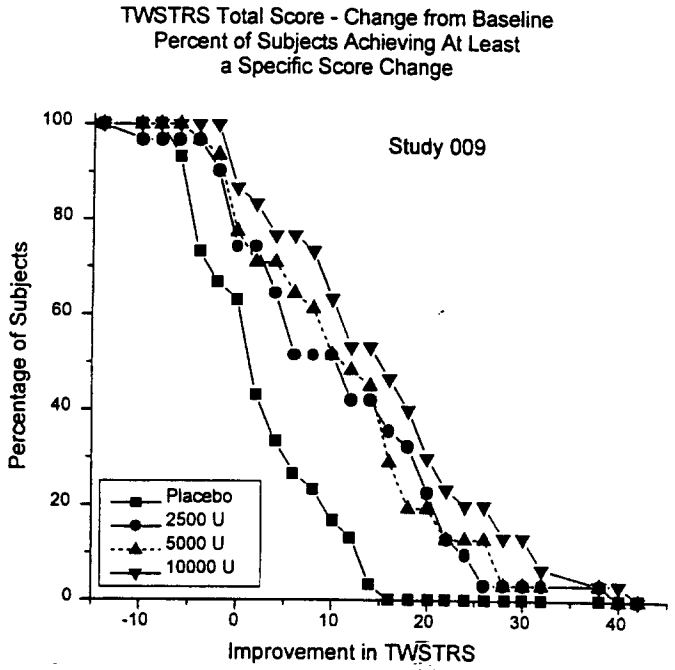
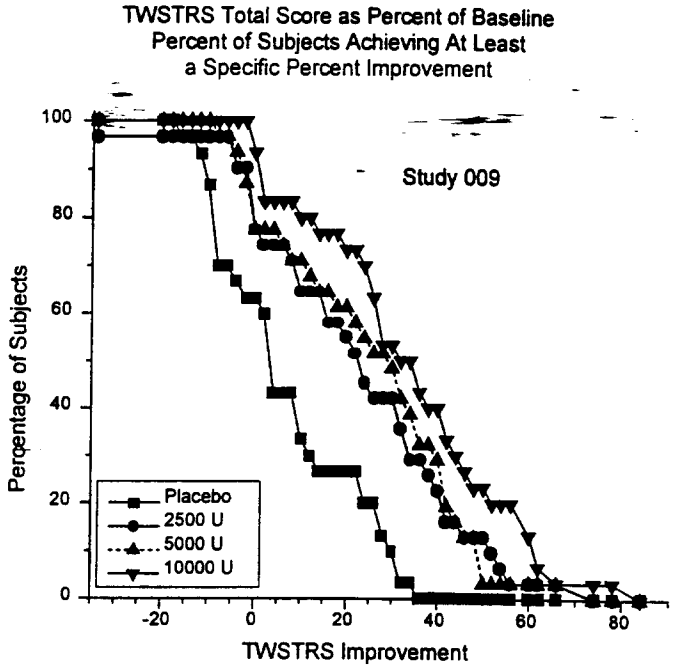


FIGURE 9



Analysis of TWSTRS Subscales

The TWSTRS subscales suggest that all three subscales contributed to the overall treatment effect.

Table 25 TWSTRS Subscale Scores - Study 009

Subscale	Time Point	Placebo n = 30 enrl	2500 U n= 31 enrl	5000 U n = 31 enrl	10000 U n = 30 enrl
Severity	Baseline	(30) 20.3	(31) 18.9	(31) 19.7	(30) 19.4
	Week 2	(30) 18.2	(30) 15.4	(30) 15.4	(30) 14.4
	Week 4	(30) 18.7	(31) 15.5 (0.051)	(31) 15.2 (0.002)	(30) 14.8 (0.0007)
Disability	Baseline	(30) 14.2	(31) 15.8	(31) 15.1	(30) 15.7
	Week 2	(30) 14.3	(30) 12.4	(30) 12.1	(30) 11.0
	Week 4	(30) 13.5	(31) 12.0 (0.017)	(31) 11.5 (0.012)	(30) 10.3 (o.0001)
Pain	Baseline	(30) 11.0	(31) 10.9	(31) 10.5	(30) 12.4
	Week 2	(30) 10.0	(30) 7.1	(30) 8.1	(30) 6.8
	Week 4	(30) 10.1	(31) 6.5 (0.002)	(31) 6.1 (0.0013)	(30) 6.0 (0.0001)

All Week 4 overall and pairwise to placebo comparisons provide $p < 0.05$,
Except Severity Subscale 2500 U vs placebo comparison, where $p = 0.051$.

Videotape TWSTRS Evaluations

The videotape TWSTRS were scored by a blinded videotape reader. The majority of videotapes were storable. In 25 (of 244 tapes) the patient visit was sufficient to score the majority of TWSTRS, but not all of the score. For these 25 subjects, the missing videotape data score was imputed with the actual clinical exam score for that visit of the subject.

Results showed that video scores at baseline were substantially less than in person scoring, ranging 19.6 to 21.3 for the baseline (vs 45-48 mean score for in person scoring). Week 4 videotape scores were not statistically significant overall, nor were the pairwise comparisons. The videotape scoring process was determined to be inaccurate and less sensitive to changes. This led to discontinuing the videotape evaluation method for subsequent studies.

EFFICACY RESULTS: OTHER EFFICACY ENDPOINTS

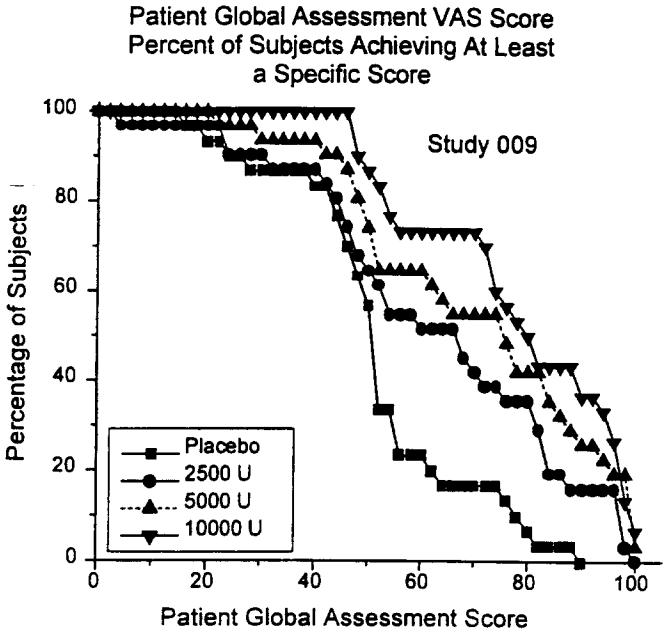
The other efficacy endpoints at week 4 also indicated that subjects had received benefit from the toxin injections in all treatment groups. The Patient Pain VAS was statistically significant at week 4 with $p = 0.005$ for the overall test, as well as statistical significance for each of the pairwise testing. The Patient Global Assessment of Change VAS was also statistically significant for the overall comparison ($p=0.0001$) as well as each of the pairwise tests. The similar was true for the Investigator Global Assessment VAS, $p = 0.0001$ overall.

Table 26 Investigator Global Assessment Mean Scores - Study 009

Endpoint	Time Point	Placebo n = 30 enrl	2500 U n= 31 enrl	5000 U n = 31 enrl	10000 U n = 30 enrl
Subjects with Week 4 TWSTRS Total 20% Improvement	n	8	18	19	23
	%	27	58	61	77
Patient Pain VAS	Baseline	52.0	45.8	49.7	46.2
	Week 2	55.1	66.9	65.1	71.1
	Week 4	52.5	67.5	70.2	75.1
	Week 4 - Change frm Basln	0.5	21.7 (0.015)	20.5 (0.008)	28.9 (0.0007)
Patient Global Assessment VAS	Week 2	51.6	64.6	62.4	72.3
	Week 4	50.6	64.3 (0.016)	69.8 (0.008)	76.6 (0.0007)
Investigator Global Assessment VAS	Week 2	53.7	65.2	64.7	70.7
	Week 4	50.2	64.4 (0.002)	69.3 (0.0001)	67.7 (0.0001)

A graph of the cumulative percent of subjects achieving specific scores provides a clearer impression of the beneficial effect of this treatment, as well as the comparisons between doses. The curves suggest a modest, but present, additional benefit to the 10000 U dose over the 5000 U dose, but even the lowest dose provides the majority of the benefit achieved with the 10000 U dose.

FIGURE 10



SAFETY RESULTS

Deaths, Serious Adverse Events and Withdrawals

There were no deaths during Study 009. There were 2 Serious Adverse events during this study, one basal cell skin cancer, and one subject with known coronary artery disease had an elective coronary catheterization performed after the routine annual exam which led to cardiac angiography to further assess cardiac disease.

There were no withdrawals from the study due to adverse events.

Incidence of Adverse Events in General

The most prominently increased adverse events associated with toxin injection was again seen to be dysphagia and dry mouth. This dose ranging study clearly suggested a dose-related increase of these events. Infection (otherwise unspecified) was also associated with toxin injection, but this was not clearly dose related, and had not been seen in the study results of Studies 301 or 302.

Injection site pain again appeared to have a modest association with toxin injection. The rates at which the previously discussed AES occurred were not markedly different in this study.

Table 27 Percentage Incidence of Adverse Events - Study 009
Events with at least 10% Incidence in Any Group

Adverse Event	Placebo n = 30 enrl	2500 U n= 31 enrl	5000 U n = 31 enrl	10000 U n = 30 enrl
Injection Site Pain	10	16	19	17
Infection (unspecified)	0	13	13	17
Headache	7	10	6	13
Pain	7	6	10	10
Flu Syndrome	0	6	10	10
Back Pain	0	3	0	10
Pain related to CD	13	10	10	23
Dry Mouth	3	3	10	33
Dysphagia	0	16	10	27
Nausea	3	10	3	10
Cough increased or Rhinitis	0	3	3	13

Note that each 3% is just 1 subject in any group; 10% is just 3 subjects

Dysphagia

All dysphagia was rated mild in intensity. No other details were provided.

PHASE 2 STUDY AN072-008

OVERVIEW

Title: A double blind placebo controlled, single dose, dose finding, safety, tolerability, and preliminary efficacy study of BotB in patients with idiopathic cervical dystonia.

This was the first of the sizable, dose ranging studies to be conducted by Athena, and provided the basis for design of the three (previously described) main studies providing evidence of safety and efficacy. However, study 008 employed only doses substantially lower than those proposed for inclusion in the product labeling, and it does not provide powerful evidence in favor of efficacy with BotTx-B. Therefore, only an abbreviated summary of design and results are provided in this review document.

CLINICAL STUDY DESIGN

Subjects met essentially the same eligibility criteria employed in studies 009, 301, and 302. Importantly, all subjects were to have been known responders to Type A botulinum toxin, and no more than 50% of subjects at any one site could be secondary non-responders to Type A toxin. Subjects were randomized in a double blind manner to 4 groups: placebo, 400 U, 1200 U or 2400 U of BotTx-B. Subjects received the study drug injections in a single treatment session, to 2 to 4 involved muscles and had telephone follow-up at 1 week, and clinic visits at weeks 2 and 4. Subjects who showed an improvement in their symptoms would continue to return at 4 week intervals until the response had declined by half, or up to 6 months. The primary efficacy endpoint was prospectively stated as the pairwise comparison of the Percentage Responders on the TWSTRS Severity Subscale at Week 4 in the middle dose group, 1200 U to placebo. Note that this study focussed upon the Severity Subscale of TWSTRS as the primary evidence of efficacy. Based on results of Study 008, the subsequent studies 009, 301 and 302 employed the full TWSTRS Total as the primary endpoint.

STUDY PERFORMANCE AND SUBJECT DISPOSITION

This study was conducted during July 1994 to February 1995. There were 85 subjects enrolled into the study (21 placebo, 21 400 U, 22 1200 U, and 21 to 2400 U) at 11 sites. There were some sites that were very sparse in total numbers and numbers per group of subjects. Only a minority of subjects completed all possible evaluation visits (7 placebo, 6 in 400 U, 6 in 1200 U, and 11 in 2400 U); most terminated the study earlier due to no apparent response or waning of response. Both circumstances were as planned by the study protocol for most subjects; however there were some subjects who were either inappropriately evaluated at 1 or 2 follow-up visits or not evaluated according to the protocol design rules. This was due to investigator errors in application of the follow-up rules, and does not affect the study data through Week 4.

Subject demographics and baseline characteristics were similar to that of the previously described studies (009, 301, 302).

PRIMARY EFFICACY ENDPOINT RESULTS AND TWSTRS OUTCOMES

The Primary Endpoint of Percent Responders at Week 4 on TWSTRS Severity Subscale in the 1200 U group was not statistically significant. There were numerous secondary endpoint analyses planned (percent responders on each subscale and total TWSTRS, subscale and total TWSTRS scores at Week 4, patient and investigator global assessments, and patient pain VAS scores; all compared pairwise of each group to placebo). Overall, there was no persuasive evidence of efficacy for either the 400 U group or the 1200 U group. However, the pairwise analyses of the 2400 U group to placebo suggested that this dose level had provided efficacy in the Week 4 evaluations. The TWSTRS Severity Subscale was not statistically significant overall, the Disability Subscale and the Total were trending (p=0.1), and the Pain Subscale was significant at p=0.03 . Additionally, the overall tests on the patient and investigator global assessments indicated a statistically significant treatment effect (p=0.007, 0.03 respectively). Due to the progressive discontinuation of follow-up in subjects after Week 4, the post Week 4 data is difficult to interpret or draw conclusions from.

Table 28: Responder Analysis at Week 4 Evaluation - Study 008

Parameter	Placebo n = 21	400 U n=21	1200 U n = 22	2400 U n = 21
TWSTRS Severity	6 (29%)	6 (29%)	7 (32%)	11 (52%) (ns)
TWSTRS Disability	4 (19%)	4 (19%)	3 (14%)	6 (29%) (NS)
TWSTRS Pain	4 (19%)	5 (24%)	5 (23%)	11 (52%) (0.025)
TWSTRS Total	3 (14%)	4 (19%)	6 (27%)	12 (57%) (0.005)

Table 29: TWSTRS Total and SubScale Results - Study 008					
Parameter	Time Point	Placebo n = 21	400 U n=21	1200 U n = 22	2400 U n = 21
Severity Subscale	Baseline	18.6	19.4	20.4	18.9
	Week 4	17	17.7	17.5	15.3
	Wk 4 Imprv. fr Basln	1.6	1.7	2.8	3.6 (0.2)
Disability Subscale	Baseline	13.1	15.6	16.0	14.1
	Week 4	13.0	14.7	15.2	12.2 (0.11)
Pain Subscale	Baseline	10.2	11.7	12.0	9.4
	Week 4	10.0	9.9	10.7	6.3 (0.005)
Total Score	Baseline	42.0	46.7	48.4	42.4
	Week 4	40.0	42.3	43.4	33.9
	Week 4 - Chg fr Bsln	2.0	4.4	5.0	8.5 (0.016)

OTHER EFFICACY ENDPOINTS

The other efficacy endpoints also provided supportive evidence that the 2400 U dose had provided some benefit to the subjects.

Table 30: Additional Efficacy Outcome Results - Study 008					
Parameter	Time Point	Placebo n = 21	400 U n=21	1200 U n = 22	2400 U n = 21
Patient Global Assessment	Week 4 (overall p=0.007)	17.4	22.0	26.3	49.5 (0.0015)
Investigator Global Assessment	Week 4 (overall p=.03)	14.5	19.6	25.7	39.7 (0.005)
Patient Pain VAS	Baseline	45.6	55.0	53.6	43.5
	Week 4	38.4	45.2	44.4	26.9
	Wk 4 - Chg fr Bsln (overall p=0.3)	7.2	9.8	9.2	16.6 (0.17)

SAFETY RESULTS

Safety evaluations did not contribute any adverse event findings not already discussed in the context of the previously reviewed studies 009, 301 or 302 (but actually conducted after this study). There were no deaths, and no serious AE. There was one withdrawal due to an adverse event. This was in the placebo group where a subject had an increase in shoulder pain and head tremor during the study, and withdrew at day 51 of the study (which was after the week 4 assessment). There was no case of dysphagia in the placebo group, and one subject with dysphagia in each of the toxin groups.

OPEN LABEL SAFETY STUDY AN072-351

DESIGN

Study AN072-351 (hereafter referred to as Study 351) became the primary open label treatment extension study for treatment of patients with cervical dystonia after completion of other IND studies.

Title: An open label, safety study of NeuroBloc (Botulinum toxin Type B) in patients with cervical dystonia.

This protocol design was initially finalized in July 1997, with a minor revision in August 1997, shortly before initiation of the study.

Objective: The evaluation of the safety of repeated doses of NeuroBloc.

Overview

This was a multicenter study in the US, Canada and UK. This open label extension study for was planned for approximately 600 subjects with CD who had participated in and completed a previous NeuroBloc study (any of the 00x studies, or Studies 301, 302 or 352). A limited number of toxin naive subjects or subjects otherwise ineligible for the prior studies would also be permitted to enroll at each of the study sites. Subjects were to receive BotTx-B injections on a repetitive basis as needed, but with a minimum interval of 12 weeks between treatment sessions. Subjects were to be evaluated at Week 4 and immediately prior to a subsequent injection, but in between only as per the investigator's usual medical practice, and without formal recorded evaluations except for collection of AES if any occur. All investigators were to obtain local IRB approval and obtain written informed consent from all subjects enrolled into the study.

Eligibility Criteria

Inclusion Criteria

History of CD for at least 1 year
TWSTRS Total of >= 20
Participation in a previous Athena BotTx-B study (# 001, 002, 003, 008, 009, 301, 302, or 352) OR each site is permitted two non-extension subjects; one toxin naive, one post surgical or phenol treatment for CD.
Male or Female, at least 18 yo

Exclusion Criteria

Receipt of any botulinum toxin injection within prior 12 weeks
Neck contractures or cervical spine disease resulting in decreased neck ROM
History of significant drug related AE in the prior BotTx-B study participation
Women who are pregnant, breast feeding, not using adequate contraception
Acute or chronic medical conditions precluding participation
Patient known to have never responded to Type A toxin
History of known other neuromuscular disease

Study Treatment

Standard vials, as for Studies 301, 302 were supplied. Subjects were assigned a new Study ID# upon enrolling in this study. Vials were packaged to contain 2500 U, 5000 U or 10000 U of toxin, and openly labeled as to content. Up to 2 vials might be used in each treatment session to make up the total dose. Toxin concentration was constant in all vials, at 5000 U/ml. Investigators selected muscles for injection and used 1 to 5 needle passes per injection muscle.

Doses between 5000 U and 15000 U, in dose level increments of 2500 U were permitted. Toxin naive subjects received 5000 U as the first dose, other subjects started at 10000 U, except for subjects who had received 15000 U in Study 352, who were permitted to continue at 15000 U.

At subsequent injection sessions the investigator was permitted to vary the dose as deemed appropriate, including dose escalation up to 15000 U.

At the time of the original application, all toxin was from manufactured at

Concomitant medications: All medications were to be documented, but none were prohibited.

Subject Evaluations

Subjects were to have history, physical exam, neurological history update and exam, clinical laboratories (including serum for antibody testing) at screening. Subjects rolling over immediately from the previous study were to have many of these assessments performed as part of the termination visit for the prior study as well, so that the visit often served as both a termination visit of the prior study and screening visit of this study. Vital signs, recording of adverse events, concomitant medications were performed at all subsequent visits. No intercurrent planned contacts with the subject were in the protocol, nor was an Investigator Global Assessment obtained.

In addition, at the Day 1 visit (day of treatment session) and the Week 4 visit, the Patient Pain VAS, the Patient Global Assessment of Change VAS, and TWSTRS score were obtained. No subject visits were planned after Week 4 except either a subsequent retreatment or termination.

Retreatment Visits could occur at least 12 weeks after the prior treatment visit, or at any time thereafter that the investigator felt appropriate. This visit became the Day 1 visit of the succeeding Treatment Cycle, with the same evaluations as noted. In addition, at retreatment visits, clinical laboratory assessments were repeated, along with serum for antibody determination.

Analytic Plan

Descriptive statistics would be used to present the assessments of safety and effectiveness. No formal hypothesis testing was planned. No explanation of the basis for the sample size of 600 was stated.

Protocol Modifications

There were no significant protocol modifications through the April 1998 period which was the interim cutoff for data to be submitted in the original BLA.

STUDY PERFORMANCE AND SUBJECT CHARACTERISTICS

Subject Disposition

The information submitted from this study in the initial filing was all information in the database in a fully quality assured status with a cut off of April 1998. The study was initiated in September 1997 and remained ongoing at the April 1998 cutoff. There were 29 sites participating in the study in North America (1 site in Canada) or UK (4 sites).

There were 260 study subjects eligible for inclusion in the submission at the time of BLA filing. Of these, 94 had received a second treatment with at least the 4 week assessment included.

Most subjects were rolled over from Studies 301, 302, 352 or earlier studies of BotTx-B. There were a total of 17 subjects who had not received BotTx-B in prior studies; 4 were toxin-naive subjects, 8 were post-neurolytic procedure, and 5 had been screened for Study 302 and failed on the F-TAT to prove Type-A resistance.

There were 8 subjects who discontinued from the study, 4 related to an AE (see safety results below), 2 for a lack of satisfactory effect, 1 for a protocol violation of receiving BOTOX treatment by a non-study physician, and 1 for unknown reason.

Protocol Deviations and Errors

There were 35 subjects who received the improper dose. Most received somewhat less than the planned dose, but details on the degree of error were not stated.

There were 25 cases with study visits outside of the proper time window, but only 4 that were more than 1 week off. There were "approximately" 25 cases of failure to collect a sample for laboratory analysis, error in injection procedure, or minor errors in consent from signing process.

For analytic purposes, subjects were categorized by dose level coarsely.

Subject Characteristics

The mean subject age was 56 years, 66% of subjects were women, 95% caucasian, mean height 168 cm, and mean weight 74.6 kg. Baseline TWSTRS for the first injection cycle (entry to the study) was 48.3.

RESULTS: DYSTONIA STATUS OUTCOME ASSESSMENT

The Dystonia status assessments at Week 4 during the study are shown in the following table.

Table 31: Outcome Assessments in Study 351

Assessment	Timepoint	Cycle 1 All Subj. n=260	Subjects with Cycle 2 Injection n = 94	
			Cycle 1 Results	Cycle 2 Results
TWSTRS Total	Baseline	48.3	49.5	46.1
	Week 4 .	35.7	37.6	38.1
	Change fr Bl	12.6	12.0	8.0
TWSTRS Severity Subscale	Baseline	21.0	21.6	20.5
	Week 4	16.3	17.0	17.1
	Change fr Bl	4.7	4.6	3.4
TWSTRS Disability Subscale	Baseline	15.6	16.3	15.3
	Week 4	12.3	13.1	13.3
	Change fr Bl	3.3	3.2	2.0
TWSTRS Pain Subscale	Baseline	11.7	11.6	10.3
	Week 4	7.1	7.4	7.7
	Change fr Bl	4.6	4.2	2.6
Patient Global Assessment	Week 4	66.1	64.9	64.6
Patient Pain VAS	Baseline	36.5	35.2	41.2
	Week 4	64.5	64.5	61.1
	Change fr Bl	28.0	28.3	19.9
Percent with 20% TWSTRS Improvmt	Week 4	57.	51	40

These results indicated continued benefit to subjects from toxin treatments. Because these are open label treatments, percentage of subjects with response or amplitude of response cannot be directly compared to the controlled and blinded studies, but in general the results are similar. The report notes that for those with two treatment sessions, there was not as large an effect observed at the second treatment session as at the first. Elan notes that the Week 4 TWSTRS outcome absolute scores and subscales are similar between cycle 1 and 2, but the pre injection score was less severe for cycle 2 than for cycle 1. Possibly there may have been incomplete washout of the first cycle before injection for the second. Consistent with this is that the Patient Global VAS was similar between Cycle 1 and Cycle 2 for these subjects.

This analysis does not describe the dose received at each of the two treatment sessions, and does not consider whether the dose was increased, remained the same, or decreased at the second cycle.

SAFETY RESULTS

Deaths, Serious Adverse Events and Withdrawals due to AES

There were 2 deaths during this study through the interim cutoff. One was an 82 yo woman who had a subarachnoid hemorrhage on day 6 after the cycle 1 injection session and died 2 days later. The second was a 71 yo woman in whom a primary lung cancer with cerebral metastases was diagnosed approximately 4 weeks after injection on cycle 1, and died shortly thereafter.

There were 11 other subjects with serious AES, in whom the SAE was not fatal. There were 8 cancers; 1 colon, 1 prostate, 4 skin cancer (1 of which was a melanoma), and 2 breast cancer. There was also one subject each with serious viral gastroenteritis, hiatal hernia with reflux, and coronary artery disease. The subject with colon cancer also had a serious intestinal obstruction, and the subject with prostate cancer an abdominal aorta aneurysm.

In addition to the two subjects who died (and are thus a form of withdrawal due to AE), there were two subjects with non serious AE who withdrew due to the AE. One subject withdrew due to persistent elevated liver enzymes (which had been elevated at baseline), and one due to toxin causally related dry mouth.

Adverse Events In General

The most frequent adverse events seen in this study are those already highlighted from experience with the prior studies. The incidence of the frequent adverse events is shown in the following table.

Table 32: Percent Incidence of Adverse Events Study 351 by Dose Cycle (at least 5% in one dosing cycle)

Adverse Event	Cycle 1 n = 260	Cycle 2 n = 94
Infection (unspecified)	7	2
Pain due to CD	6	3
Dysphagia	17	13
Dyspepsia	5	5
Dry Mouth	36	22
Headache	5	2

Comment:

Dose levels across the treatment cycles were not analyzed, nor was incidence provided for constant-patient subsets.

The incidence of dysphagia is somewhat lower in this study than in the previous studies at the 10000 U dose level. The reasons for this are unknown, but include the possibility of

lower doses used in some patients, a different distribution of dose amongst muscles compared to the subject's dose pattern in their prior study, other forms of investigator learning effect with individual subjects, and selection of subjects (self or investigator) for the entry into the extension study. Decreased sensitivity to AES between the controlled and open label study remains a possibility as well.

The frequent adverse events were also examined by distribution of severity, and are shown below for the AES of dysphagia and dry mouth, the only events with an aggregate incidence of at least 10% within one dosing cycle.

Table 33: Percent Incidence of Frequent AE in Study 351 by Severity (at least 10% total in one dosing cycle)

AE	Dose cycle	Mild	Moderate	Severe
Dysphagia	Cycle 1	11.5	4.2	0.4
	Cycle 2	7.4	3.2	1.1
Dry Mouth	Cycle 1	20.4	10.0	5.4
	Cycle 2	13.8	3.2	3.2

There was 1 case with severity unspecified for each row, except for Dry Mouth in Cycle 2, where 2 cases of unspecified severity

Antibody Formation

There were 236 subjects with at least one serum sample tested contained in the interim report. Of these, 34 subjects (14%) had a positive ELISA titer. The summary report did not identify how many serum samples were tested at each of the timepoints (screening, at day of injection session 2, at day of injection session 3 timepoints), so that actual rates of positive ELISA are not known. One serum sample of the ELISA positive samples did show the presence of neutralizing antibodies in the mouse protection assay, with survival of 4 of 4 tested mice, repeated and again showing 4 surviving mice. The titer level of neutralizing antibodies was not determined due to assay difficulties.

OPEN LABEL SAFETY STUDY AN072-352

DESIGN

Title: An open label, dose-escalation safety and tolerability study of Botulinum toxin type B (BotB) in patients with cervical dystonia.

This protocol was finalized in March 1997, and amended in May 1997 and July 1997 before the study actually commenced.

Objective: To evaluate the safety of BotTx-B in subjects not previously exposed to BotTx-B of up to 15000 U, over a three session repeat dosing schedule.

Overview

Study AN072-352 (referred to as Study 352) was an open label study with three treatment sessions per subject designed to assess the safety of increasing the dose of BotTx-B beyond the 10000 U for which there was already clinical experience.

This was a multicenter, open label, non-randomized, intra subject dose escalation study for subjects who had not previously been exposed to BotTx-B. Subjects who completed this study would be offered enrollment into Study 351. Up to 160 subjects were to be enrolled, in approximately 18 centers, in both North America and UK.

Subjects were classified as Type-A responsive or (secondary) resistant. Subjects who had previously had a worthwhile response to Type-A toxin, then had lost such responses, including to increased dose of BOTOX, and had Type A resistance shown on an F-TAT were classed as Type-A resistant.

Centers were permitted to have no more than 50% Type-A resistant subjects enrolled at any time. Centers were not required to enroll A-Resistant subjects. IRB approval was to be obtained by all sites, and informed consent from all subjects.

Eligibility Criteria

Inclusion Criteria

History of CD at least 1 year duration
At least two eligible muscles are involved in the CD.
Levator scapulae, Scalenus medius, anterior.
Semispinalis capitus, Splenius captus; Sternocleidomastoid, Trapezius.
CD of at least moderate severity, defined as
TWSTRS Total >= 20 ; Severity >= 10 ; Disability >=3; Pain >=1
Male or Female, at least 18 yo, body weight at least 46kg.
A history of perceived beneficial response to previous use of Botulinum Toxin Type A

Exclusion Criteria

Received botulinum toxin within prior 4 months
Neck contractures of cervical spine disease producing decreased ROM
Women who are pregnant, breast feeding, inadequate contraception.
Other medical disorders making patient unsuitable
Tetanus toxoid within prior 4 months

Study Treatment

Standard formulation vials, as were used in Studies 301 & 302 were supplied. Vial contents were 2500, 5000, or 10000 U, at 5000 U/ml as previously, and openly labeled as to content. All toxin used in this study was manufactured in the [redacted] facility.

Subjects were administered 10000 U on the first treatment session. After the subject returned to their baseline CD status, they were eligible for a repeat injection, which could be at the next higher dose level (12500 U). This could be repeated again for a third cycle of treatment, so that subjects were eligible for a maximum of three treatment cycles, at maximum doses of 10000 U, 12500 U, and 15000 U on each successive cycle. Subjects who developed AE due to toxin would not dose escalate, and would terminate the study at that time, without having completed all three treatment sessions.

Concomitant medications: All medications were to be documented, but none were prohibited.

Subject Evaluations

Subjects underwent history, physical exam, specialized neurological exam (including TWSTRS) and history, and clinical chemistries (including serum for antibodies) at screening.

A week 1 follow-up phone call for adverse events was employed. Follow-up visits included vital signs, adverse event and concomitant medication recording. Each cycle employed the same sets of evaluations at each of the designated visits within that cycle. Disease specific evaluations consisted of:

Day 1 visit (treatment visit for each cycle)
Patient Pain VAS
Patient Global VAS Assessment
TWSTRS

Week 2 Visit
TWSTRS

Week 4 Visit
Patient Global VAS
Patient Pain VAS
TWSTRS
Investigator Global VAS Assessment

Weeks 8, 12, 16 Visits
- Patient Global VAS
- Patient Pain VAS
- TWSTRS
- Investigator Global Assessment

Any of the Week 4, 8, 12, 16 visits could also become the Day 1 visit for the next treatment cycle, if the subject has returned to baseline status.

Day 1 visit of 12500, and 15000 dose cycles also include a re-check of clinical laboratory values, and included a serum sample for Antibody determination.

The termination visit also include clinical chemistries, physical exam, serum for antibodies.

Analytic Plan

Only descriptive statistics for the data collected were planned. No formal hypothesis testing was to be conducted. Sample size of 160 subjects was selected, and stated as not based statistical considerations. The basis for sample size selection was not stated.

STUDY PERFORMANCE AND SUBJECT CHARACTERISTICS

Subject Disposition

The study was initiated in June 1997. Information contained in the initial BLA submission was information that had at least the Week 4 visit entered into the database and in a fully quality assured status by the date of the cutoff of April 1998. The study remained ongoing at that time.

At the time of the interim data cutoff, there were 138 subjects enrolled with at least a Week 4 visit of cycle 1 in the database. These were enrolled across 16 different sites. All subjects had been treated in the 10000 U cycle, 103 had been treated at 12500 U cycle, and 55 at 15000 U. Thus there were 35 subjects who had participated only at the 10000 U cycle, 48 subjects only through the 12500 U cycle.

There were 17 subjects who withdrew permanently, 7 for adverse events, 5 classified as due to non-AE related patient request (4 for lack of efficacy and 1 for being unable to participate due to incarceration). There were also 5 patients that withdrew for other reasons including travel costs, moving out of clinic area, and preferring resumption of treatment with BOTOX.

Comment:

The differentiation of "patient request" between the first described set of 5 and the second is not clear.

At the time of the interim cutoff there were only 12 subjects who had fully completed the 3 cycle course; 109 subjects were still enrolled and continuing the study at this time.

Protocol Conduct Violations

There were 3 subjects with eligibility violations. There was 1 each of a history of psychiatric disorder, non-compliance with the contraception requirement, and receipt of tetanus toxin within 4 months of study enrollment.

Other forms of protocol violations include 90 visits outside of the specified time window, and 29 cases of scheduled evaluations not completed or not completed correctly.

Treatment Characteristics

Comment:

As for other studies, no summarization regarding actual doses administered, number of muscles, or which muscles were injected was supplied by the Elan.

Protocol Modifications

There were no significant protocol modifications made after the study commenced.

Study Population Characteristics

The mean age of subjects was 53 years, 64% of subjects were female, and 95% of subjects were caucasian. Mean height is 169 cm, and weight 73 kg. Of the subjects enrolled, 78% were type-A responsive. The mean baseline TWSTRS Total score was 47.0 for the initial treatment cycle. In general, these subjects were similar to those enrolled in the previously reviewed studies.

RESULTS: DYSTONIA STATUS ASSESSMENTS ON STUDY

There were substantially fewer subjects at week 8 visits than treated for the cycle. This was not differentiated between simply interim database cutoff or subjects who had already discontinued follow-ups for that cycle by having previously returned to baseline and proceeded on to the subsequent cycle's treatment session.

Comment:

Also not presented by Elan was comparison of response on different dose levels relying on the same patient subset for each dose level. However, the data presented suggest that there is no advantage in amount of benefit to injection with higher doses of toxin beyond 10000 U.

Table 34: Dystonia Assessments on Study - Study 352				
Assessment	Time Point	10000 U Cycle n= 138 treated	12500 U Cycle n= 103 treated	15000 U Cycle n = 55 treated
TWSTRS Total	Cycle Baseline	47.0	47.6	48.2
	Week 2	37.8	37.0	38.1
	Week 4	37.1	36.6	37.7
	Week 4 Improvement	9.9	11.0	10.4
	Week 8	41.0 (n=118)	38.2 (n=93)	38.5 (n= 28)
Patient Global VAS	Week 4	58.5	62.3	60.7
	Week 8	51.0	55.3	52.2
Investigator Global VAS	Week 4	64.6	66.7	66.6
	Week 8	58.7	62.4	65.5
Patient Pain VAS	Cycle Baseline	39.9	36.4	38.4
	Week 4	61.2	63.0	64.2
	Week 8	49.0	55.2	58.5
	Week 4 Improvement	21.2	26.6	25.8
	Week 8 Improvement	9.5	17.6	23.0
Percentage with at least 20% Improvement TWSTRS	Week 4	51%	57%	51%
TWSTRS Severity Subscale	Cycle Baseline	19.9	19.8	20.6
	Week 4	16.4	16.6	16.4
TWSTRS Disability Subscale	Cycle Baseline	15.6	16.2	16.3
	Week 4	13.0	13.1	13.3
TWSTRS Pain Subscale	Cycle Baseline	11.5	11.6	11.3
	Week 4	7.7	7.9	8.1

Note: All Week 8 evaluations fewer than the number treated, as noted for TWSTRS

SAFETY RESULTS

Deaths, Serious Adverse Events and Withdrawals due to AE

There was 1 death on study, due to non-Hodgkins Lymphoma. There were 4 additional SAE's, consisting of a psychotic depression, cholelithiasis, renal calculus, and a cellulitis of the right calf.

There were 7 subjects who withdrew due to AE. The 6 during the 10000 U dosing cycle were due to the psychotic depression, 3 for dysphagia (1 also with dyspepsia and dry mouth) and 1 for dry mouth with other sensory symptoms. The 1 withdrawal for AE at the 12500 U cycle was the subject with the SAE of non-Hodgkins lymphoma.

Adverse Events in General

The adverse events occurring most frequently during this study were similar in nature to those identified in the previously reviewed studies in this document.

Table 35: Percent Incidence of frequent AE in Study 352 by Dose Cycle (at least 10% in one dosing cycle)

AE	10000 U cycle n = 138	12500 U cycle n = 103	15000 U cycle n = 55
Infection (unspecified)	10	17	9
Injection site pain	16	17	5
Dysphagia	34	38	18
Dyspepsia	15	7	2
Dry Mouth	51	40	36

Comment:

Of note however, the incidence of dysphagia and dry mouth were significantly higher than that seen in any of the prior studies. The incidence rates across dose levels was not provided for constant-patient subsets. The decreased rate of dysphagia in the third dosing cycle cannot be interpreted without further information.

The most frequent of these AES were examined with incidence subdivided by severity. The severity of these adverse events does not appear to increase with increasing dose. However, given the small numbers of subjects with AES, and the uncertainties regarding selective drop-out from higher doses, conclusions cannot be reached based on this information.

Table 36: Percent Incidence of frequent AE in Study 352 by Severity (at least 10% total in one dosing cycle)

AE	Dose cycle	Mild	Moderate	Severe
Infection (unspecified)	10000 U	7	2	1
	12500 U	11	5	1
	15000 U	9	0	0
Injection site pain	10000 U	7	9	0
	12500 U	7	10	0
	15000 U	0	5	0
Dysphagia	10000 U	22	11	1
	12500 U	23	13	2
	15000 U	13	4	2
Dyspepsia	10000 U	7	6	2
	12500 U	3	4	0
	15000 U	0	2	0
Dry Mouth	10000 U	34	16	1
	12500 U	26	14	0
	15000 U	36	0	0

Antibody Formation

Serum samples for 115 subjects were assayed, of which 22 subjects (20%) had positive ELISA titers on at least one of the tested timepoints (screening, day of the 12500 U dose, day of the 15000 U dose, termination). Elan did not provide the number tested at each of these timepoints. There were 10 subjects with titers above 0 at screening, 14 prior to the 12500 U dose, 11 prior to the 15000 U dose, and 7 prior to the 15000 U dose. This may be suggestive of an increase with time or repeat administration, but awaits further information on the distribution of positive ELISA. None of the ELISA positive samples that could be analyzed were positive for mouse neutralizing antibodies, however not all samples were tested.

INTEGRATED SUMMARY OF SAFETY

BASIS OF THE SAFETY DATASET

The Integrated Summary of Safety (ISS) report indicates that there were a total of 255 subjects amongst all the placebo controlled studies, who received a total of 268 treatment sessions (a few subjects who had participated in early studies had been re-enrolled into later efficacy studies). There were a total of 438 subjects with 738 treatment sessions in the uncontrolled studies. Most of the subjects in the uncontrolled studies had been previously participated in the placebo controlled studies. While this is intended to be chronic therapy, there were only 156 subjects with 2 or more treatment sessions, and 64 with 3 treatment sessions in the uncontrolled safety dataset.

Combined, there were 574 total subjects in all studies (including placebo subjects), 371 in controlled studies and 203 who participated in only the uncontrolled studies. Of these, 531 subjects were treated with toxin, in 1006 total dosing sessions. There were 403 subjects who received at least 1 dose within the 7500 to 11000 U range, and 76 subjects who received a dose of 15000 U or more (most in Study 352; a few in Study 351).

For summary analyses, Elan grouped subjects across studies or treatment sessions according to dose received, in coarse categories.

ADVERSE EVENT TABULATIONS

Frequent Adverse Events in Controlled Studies

The overall adverse event rates again suggest that dry mouth and dysphagia are the most prominent associated AES, and that there is a dose related progressive increase. Dyspepsia may have a small increase, but only in the highest dose levels. Other adverse events appear to not have a strong association with toxin use, although a threshold effect that is just showing up in the highest dose level is not excluded.

Table 37 Percentage Incidence of Most Frequent Adverse Events Only Placebo Controlled Studies At Least 5% Incidence in a Group

Adverse Event	Placebo n = 123	<=2500 U n= 92	2500 - 7500 U n = 67	7500 - 10000 U n = 106
Injection Site Pain	7	10	12	15
Infection (unspecified)	13	5	19	15
Headache	9	22	16	11
Pain	9	13	7	13
Flu Syndrome	5	2	9	8
Asthenia	7	10	0	5
Neck Pain	6	8	3	3
Back Pain	2	4	4	7
Accidental Injury	4	1	4	5
Chills	3	8	0	2
Malaise	1	5	0	1
Pain related to CD	14	0	16	17
Dry Mouth	2	2	12	34
Dizziness	3	3	3	6
Torticollis	4	0	4	8
Dyspepsia	4	3	1	10
Dysphagia	2	9	10	25
Nausea	8	13	3	8
Cough Increased	4	2	6	7
Rhinitis	7	2	1	5
Pharyngitis	6	2	3	4
Bronchitis	6	0	3	0
Myasthenia	4	9	4	7
Arthralgia	4	0	0	7

Frequent Adverse Events in Controlled Studies within 4 Weeks of Treatment

Limiting examination to only those AES developing promptly after toxin treatment (within 4 weeks) to focus on treatment related events again reveals dysphagia and dry mouth the most prominent AES, with injection site pain also suggested to have an association. Dyspepsia may have an association with the highest dose, but the other AES appear to not be associated with toxin administration.

Table 38 Percentage Incidence of most frequent adverse events
Only Placebo Controlled Studies - Within First 4 Weeks Only
At Least 5% Incidence in a Group

Adverse Event	Placebo n = 123	<=2500 U n= 92	2500 - 7500 U n = 67	7500 - 10000 U n = 106
Injection Site Pain	7	10	12	15
Infection (unspecified)	7	2	3	8
Headache	7	21	12	9
Pain	7	13	4	8
Flu Syndrome	2	1	4	5
Asthenia	5	10	0	5
Neck Pain	4	7	0	0
Back Pain				
Accidental Injury				
Chills	2	8	0	2
Malaise	1	5	0	1
Pain related to CD	9	0	3	8
Dry Mouth	2	2	12	34
Dizziness	3	3	3	5
Torticollis	4	0	3	5
Dyspepsia	2	2	0	8
Dysphagia	2	9	10	25
Nausea	7	13	3	8
Cough Increased				
Rhinitis				
Pharyngitis				
Bronchitis				
Myasthenia	4	9	4	7
Arthralgia				

Blank rows did not meet 5% threshold for this table, but had been in Table 37

Frequent Adverse Events in Uncontrolled Studies Divided by Dose Received

Table 39 provides comparisons of AES occurring in subjects divided by the dose received, pooled across all studies. Because these were non-concurrent studies, not randomized, and includes variable overlap of subjects between dose levels, definitive conclusions on dose-relatedness cannot be formed from this table. However, the previously identified AES remain the most prominent. Dysphagia and dry mouth remain the AES with greatest frequency and most clear dose-related increase in incidence. Other AES, as noted before are of some concern, but not as prominent as dysphagia and dry mouth.

Table 39 Percentage Incidence of Most Frequent Adverse Events
UN-Controlled Studies Divided by Dose and All Toxin Exposed Subjects Combined
At Least 5% Incidence in a Group Or 10 Subjects

Adverse Event	2500 - 7500 U n = 37	7500 - 11000 U n= 378	11000 - 15000 U n = 152	=> 15000 U n = 76	All Toxin Subjects n = 531
Injection Site Pain	11	8	12	4	15
Infection (unspecified)	19	8	13	7	18
Headache	16	6	2	4	14
Pain	14	4	7	1	13
Flu Syndrome	5	5	4	4	9
Asthenia	14	3	3	1	8
Neck Pain					5
Back Pain					5
Accidental Injury	0	4	3	3	6
Fever					2
Chills					2
Malaise					2
Pain related to CD	3	6	7	3	12
Dry Mouth	22	42	37	29	41
Dizziness	3	3	2	3	6
Torticollis					4
Hypertonia					3
Insomnia					3
Depression					2
Dyspepsia	3	9	8	1	10
Dysphagia	8	23	32	14	29
Nausea	5	3	3	1	8
Diarrhea					4
GI Disorder (not specified)					3
Cough Increased	5	1	1	0	4
Rhinitis					4
Pharyngitis	5	4	4	1	6
Voice Alteration					3
Sinusitis					3
Bronchitis					2
Myasthenia	5	3	3	4	7
Arthralgia	5	4	7	4	6
Myalgia					3
Amblyopia	0	3	1	0	2
Rash	8	1	2	1	3

Note that criteria of 10 subjects includes incidences of 3% in 7500 - 11000 U and 2% in All toxin subjects column

This table includes multiple treatment sessions on many subjects; Subjects with multiple occurrences of same AE were counted only once in any column, but may be counted in more than 1 column

Frequent Adverse Events Subdivided by Severity

Table 40 provides a subdivision of adverse events by severity. The most important aspects of this table are that for dysphagia and dry mouth, the most important dose related adverse events, the majority of events are of mild nature, but there are some events that are rated as severe. This is higher for dry mouth than dysphagia, but balanced by the fact that severe dysphagia can be more medically risky than severe dry mouth. Nonetheless, even moderate grade events in these two categories have been responsible for patients discontinuing repetitive injections with BotTx-B. Additionally, injection site pain, which has a moderate dose relatedness, has half of these events of moderate or severe grade, making this an important issue for the tolerability of the treatment.

Table 40 Percentage Incidence and Severity Distribution of Most Frequent Adverse Events in Aggregated Toxin Exposed Subjects At Least 10 Subjects Total (2%) with AE

Adverse Event	All Toxin Subjects n = 531	Mild	Moderate	Severe
Injection Site Pain	15	7	6	1
Infection (unspecified)	18	11	6	1
Headache	14	10	3	2
Pain	13	6	5	1
Flu Syndrome	9	4	5	0
Asthenia	8	5	2	0
Neck Pain	5	2	3	1
Back Pain	5	2	2	0
Abdominal Pain	4	2	1	0
Accidental Injury	6	3	3	0
Fever	2	2	0	0
Chills	2	2	0	0
Malaise	2	1	1	0
Pain related to CD	12	4	6	2
Dry Mouth	41	24	13	4
Dizziness	6	4	2	0
Torticollis	4	2	2	1
Hypertonia	3	1	2	1
Insomnia	3	2	1	0
Depression	2	1	1	0
Dyspepsia	10	5	4	1
Dysphagia	29	19	8	1
Nausea	8	6	2	0
Diarrhea	4	3	1	0
GI Disorder (not specified)	3	2	1	0
Cough Increased	4	3	1	0
Rhinitis	4	3	1	0
Pharyngitis	6	5	2	0
Voice Alteration	3	2	1	0
Sinusitis	3	2	1	0
Bronchitis	2	0	2	0
Myasthenia	7	5	2	0
Arthralgia	6	2	3	1
Myalgia	3	1	2	0
Amblyopia	2	2	1	0
Rash	3	3	0	0

Note that criteria of 10 subjects includes incidences of 3% in 7500 - 11000 U and 2% in All toxin subjects column

This table includes multiple treatment sessions on many subjects; Subjects with multiple occurrences of same AE were counted only once in any column, but may be counted in more than 1 column

Combined Analyses of Adverse Events Associated with Specific Muscle Injection

Elan repeated their dose-tertile subdivision method to analyze for a dose-related association of toxin treatment to specific muscles. Their analyses were in the ISS were conducted with subjects from Studies 301 and 302 pooled.

The trapezius muscle appeared to have a dose-related association with dry mouth across the no-toxin / low / medium / highest dose cohorts (10%, 35%, 38%, 58%), but less uniformly for dysphagia (10%, 24%, 15%, 42%) with dose tertiles means of 1147 U, 2308 U, 4067 U (group n's of 144, 17, 13, 12), but most of these subjects were from Study 301, so this adds little weight over that prior analysis.

The SCM had larger tertile sizes (n = 29, 32, 29 [& 96 with no SCM toxin]) with mean doses of 986 U, 2008 U, 3816 U. The association appeared to occur for both dry mouth (7%, 14%, 38%, 34%) and for dysphagia (4%, 14%, 22%, 38%).

EXPLORATORY ANALYSES OF STUDY RESULTS

EFFICACY WITHIN NARROW RANGES BASELINE DISEASE SEVERITY

As noted, there are toxicities associated with this treatment. Therefore, an appropriate question to examine is whether or not benefit from this therapy is received by subjects in some identifiable, differential manner, that might imply the risk-benefit comparison is not the same for all subjects within the subsets. The subjects enrolled in these studies had a range of severities of dystonia. While in general the subjects enrolled had moderate severity dystonia, the range of TWSTRS Total score at baseline was broad. There was not a disparity between treatment groups in the randomized studies, so that differential benefit could not have produced a false appearance of overall efficacy. However, differential efficacy would be important information to provide to patients and physicians in considering if, and how, to use this therapy.

The subjects in the controlled studies were subdivided by baseline TWSTRS scores to give quartile groups (as best as achievable). Because the controlled studies had relatively few subjects in each subset when each treatment group was subdivided, studies were combined where feasible to have sufficient subjects in each subdivision to make reasonable comparisons.

This analysis are shown in the following figures for the example outcomes of Change in TWSTRS to Week 4 as a Percentage of Baseline Score, and for the Patient Global Assessment. These figures indicate that the benefit from toxin treatment is received by patients in all quartiles of baseline TWSTRS scores. There is no end of the spectrum of tested severities that does not receive benefit. The Subset 3 in left figures, subjects with baseline scores 46 to 53, appears to have not received benefit; however this is most likely artifact of the post hoc subdivision process. The Placebo group for these subsets have appears to have an inconsistently large improvement from baseline, compared to the adjacent subsets. When subjects with nearly the same range of baseline scores are examined in the three combined studies on the right-sided figures, this apparent inconsistency has disappeared and efficacy is clearly present. Thus, subjects appear to obtain benefit from the treatment irrespective of the baseline TWSTRS scores.

FIGURE 11

Change in TWSTRS as Percent of Baseline: Subsets by Baseline TWSTRS Score

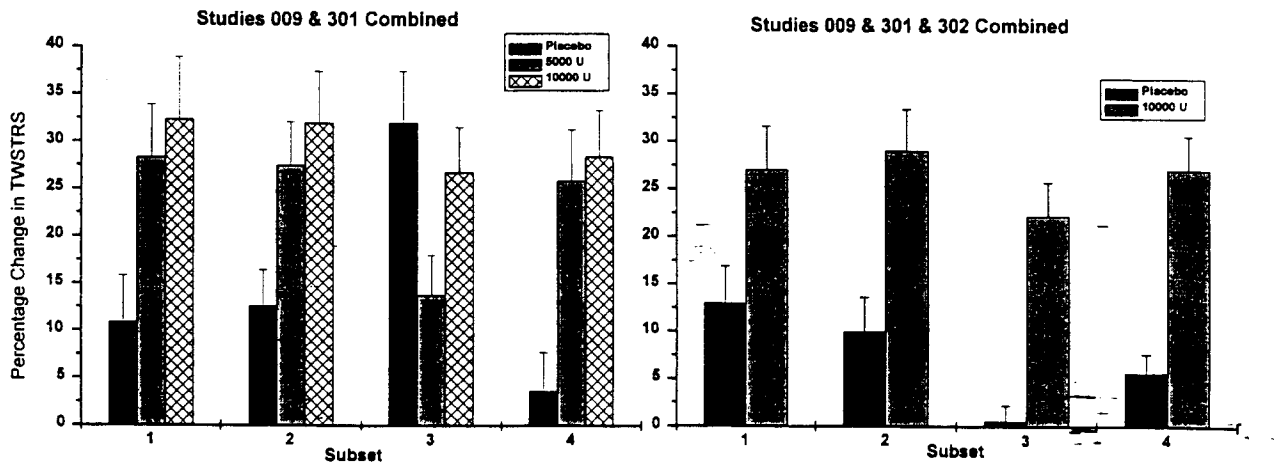


FIGURE 12

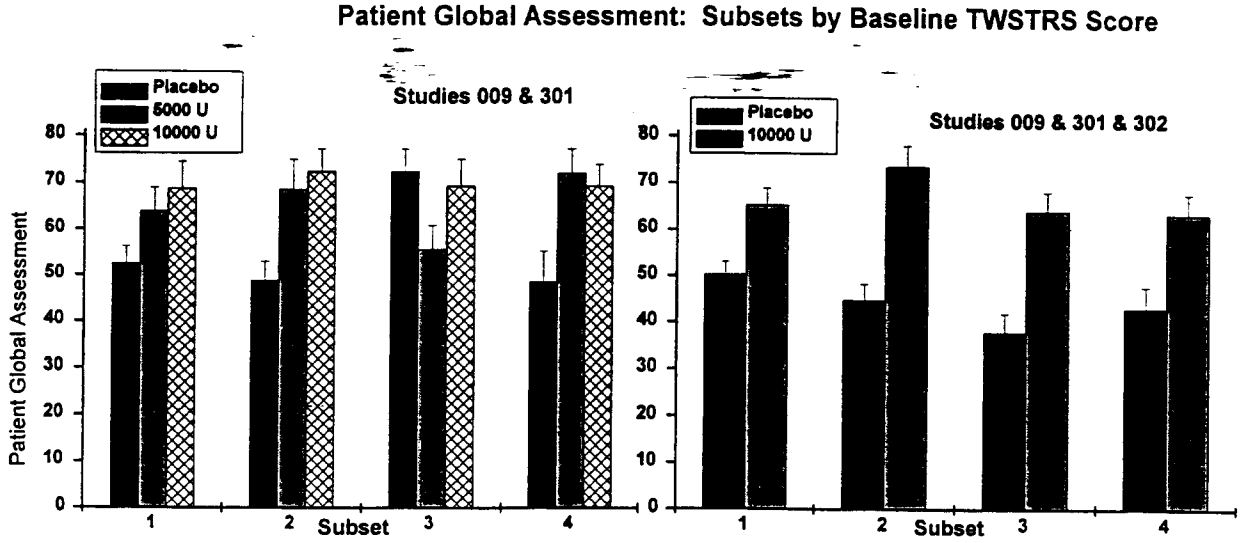


Table X: Details of Subsets by Baseline TWSTRS Score							
Subset #	Studies 301 & 009 Combined				Studies 301 & 302 & 009 Combined		
	TWSTRS Range	Placebo	5000 U	10000 U	TWSTRS Range	Placebo	10000 U
1	20 - 39	17	19	12	20 - 41	28	24
2	40 - 45	21	16	18	42 - 48	28	24
3	46 - 53	18	16	18	49 - 55	27	29
4	54 - 72	9	16	19	56 - 72	21	29

EFFICACY WITHIN SUBSETS BY SEX

Subsets of subjects were created by dividing the controlled trial subjects by sex and examining the efficacy outcomes in combined studies datasets. Both men and women appear to obtain benefit from the treatment. There is a suggestion that men obtain lesser benefit from the 5000 U dose than from the 10000 U dose, unlike the women who appear to obtain equal amounts of benefit from both doses. While the men form a minority of the patients with this disorder, and a similar minority of subjects in these studies, the subsets were sufficiently large to provide some reliability to these observations.

FIGURE 13

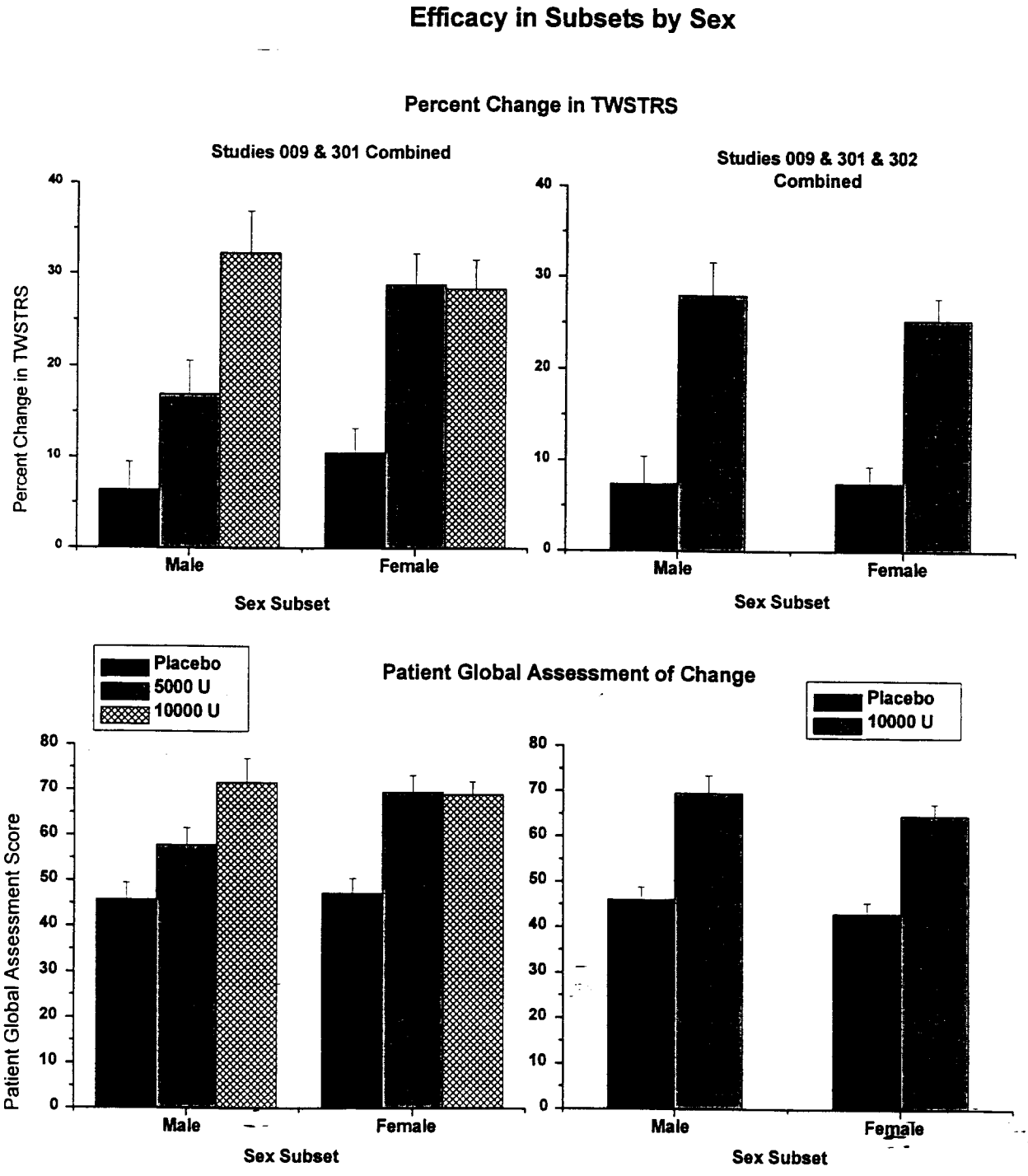


Table X: Details of Subsets by Sex					
Sex	Studies 301 & 009 Combined			Studies 301 & 302 & 009 Combined	
	Placebo	5000 U	10000 U	Placebo	10000 U
Male	24	27	21	36	33
Female	42	40	46	68	73

EFFICACY WITHIN SUBSETS BY AGE

Subjects were divided into 10 years of age range groups and efficacy outcomes examined. Subjects of all ages appear to have derived benefit.

FIGURE 14

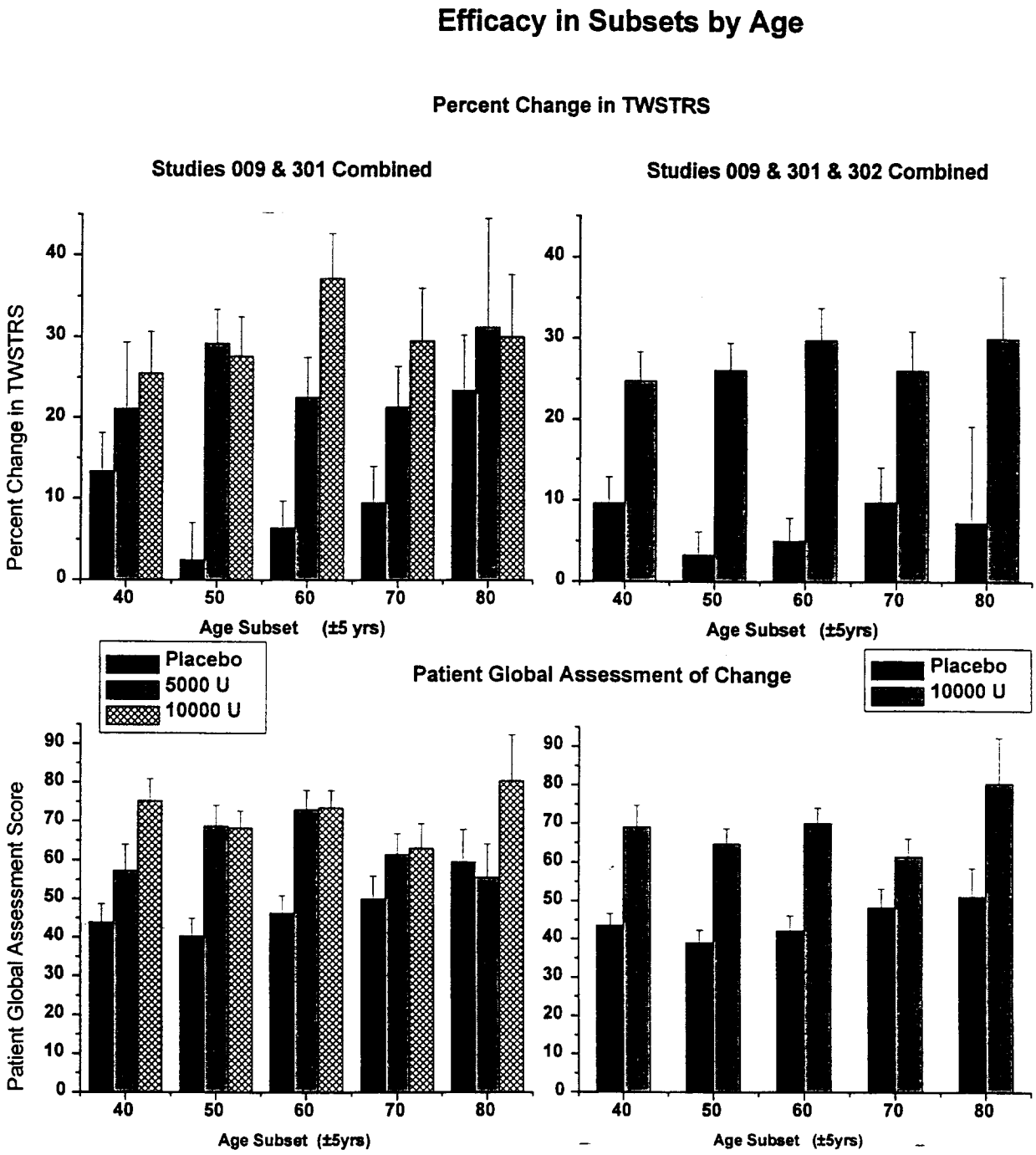


Table X: Details of Subsets by Age					
Age Range	Studies 301 & 009 Combined			Studies 301 & 302 & 009 Combined	
	Placebo	5000 U	10000 U	Placebo	10000 U
35 - 45	13	12	10	22	15
45 - 55	14	23	22	26	33
55 - 65	23	15	13	30	24
65 - 75	11	12	16	17	26
75 - 90	3	3	4	5	4

Note: Age range includes upper limit, excludes lower limit

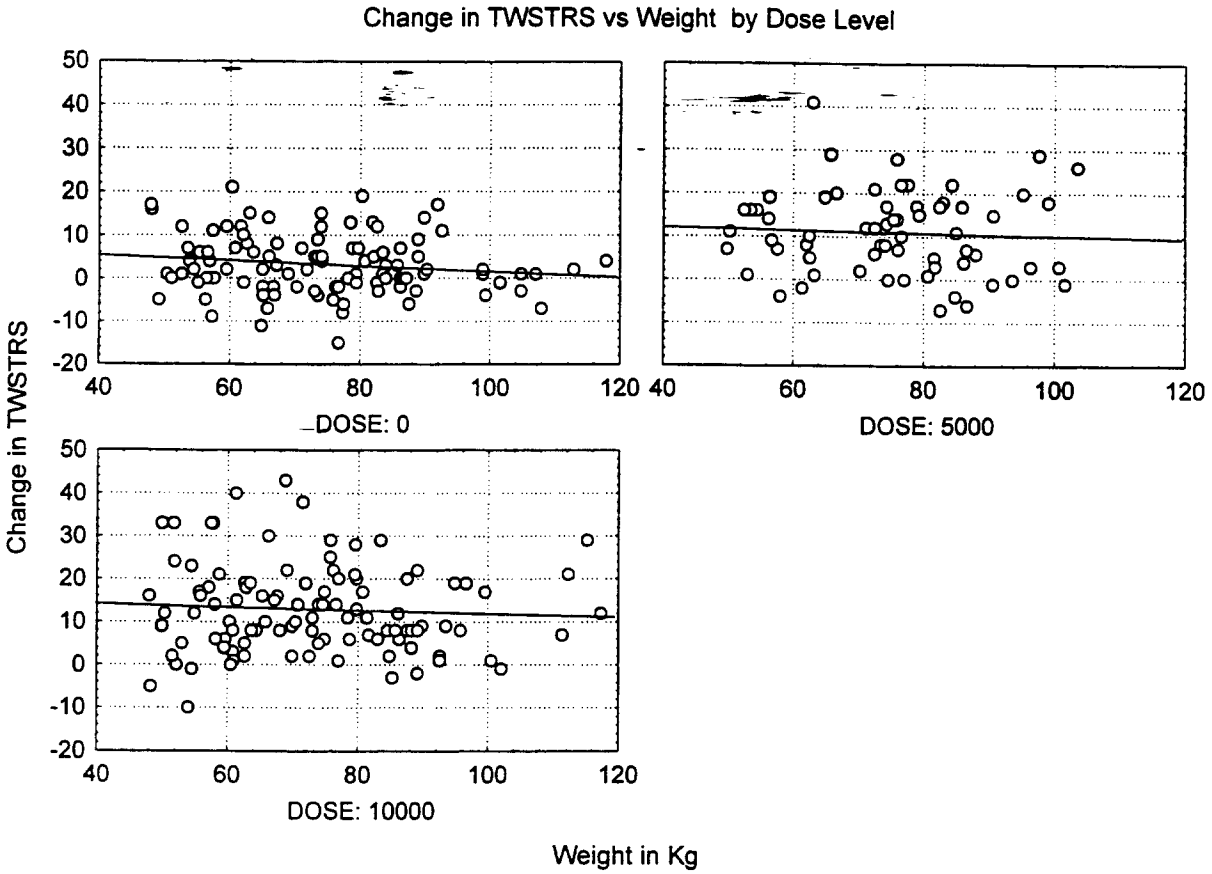
EFFICACY ASSOCIATED WITH SUBSETS BY RACE

As shown in the tables of subject demographics for each study reviewed above, there were relatively few subjects in these studies not within the caucasian subset. No ability to analyze for differential effects in subsets by race is afforded by these data.

EFFICACY RELATED TO VARIATIONS IN SUBJECT WEIGHT

There appeared to be no evidence of a differential efficacy with subject weight. Subjects across a the majority of the range of weights showed improvements in TWSTRS scores with toxin treatments, while the placebo group subjects showed a very consistent outcome behavior over the range in weight.

FIGURE 15



EFFICACY AND RISK COMPARISON BETWEEN 5000 AND 10000 U AND HIGHER DOSES

An important aspect of the labeling is recommendation of dosing. Elan has proposed recommendation of the 10000 U dose, with statements supporting the safety of using higher doses. As noted in many of the analyses presented previously in this document, there appears to be little additional benefit to a dose of 10000 U over that offered by 5000 U (see Figure 4, 5, and 7 for between dose comparison within Study 301 and Figures 8, 9 and 10 for the comparison within Study 009). This lack of dramatic additional benefit is also consistent across the range of baseline disease severity (see Figures 11 and 12 for between dose comparisons based upon pooling of data from both Study 009 and Study 301). Study 009 suggests a dose of 2400 U offers significant efficacy, and Study 009 suggests there may be only modest benefit to a dose of 5000 U over 2500 U. Although incomplete at present, the available information from Study 352 suggests there is no benefit to 15000 U over 10000 U. Taken together, this information suggests that the great majority of the benefit of the treatment is achieved by a dose of 5000 U, that at least substantial benefit is achieved by a dose of 2500 U, and that no perceived additional benefit is achieved by doses beyond 10000 U.

Assessment of duration of benefit was not a major objective of these studies. To the degree that the controlled studies were able to assess this, there was not evidence of a significant difference in duration of benefit between the 5000 U and 10000 U doses.

As a counterpoint to this is the adverse event profile. Dysphagia and dry mouth are the most prominent adverse events associated with BotTx-B use, and these have a clearly dose-related

incidence profile. The incidence of these events is at least doubled between 5000 U and 10000 U. Incidences may also be increased between 2500 U and 5000 U. Information on incidences above 10000 U is sparse at present, and not well compared with lower dose incidences. Dysphagia has the potential to be severe, and the potential to severely interfere with a patient's life. Both dysphagia and dry mouth, even without reaching the level of being a serious AE, have been severe enough to induce subjects to discontinue repeat treatments in clinical studies. Thus, these events are important to the patients, and do influence the risk to benefit comparison.

Furthermore, and additional important aspect to considered is that all studies have enrolled only subjects who have a history of perceived favorable risk-benefit to botulinum toxin type A use. Thus, these studies eliminated before enrollment any subjects who might be at excess sensitivity to botulinum toxin IM injections as a general treatment strategy, and the generalizability is somewhat impaired by this concern. Truly toxin-naive subjects may have a higher propensity for adverse events, thus forming a different risk to benefit assessment.

Therefore, recommendation of 10000 U dose for all patients is unwarranted. 5000 U as an initial dose can be expected to provide the majority of benefit at lower risk than 10000 U, and 2500 U should be considered for any subject with any neurologic or muscular impairment.

EFFICACY EFFECTS OF ANTIBODY FORMATION

The data submitted do not provide a sufficient basis to assess either the long term potential for antibody formation, nor the long term consequences of antibody formation. Relatively few subjects had ELISA positive serum in these studies either at start of the study or at the follow up after a single treatment session. Of these, only one serum sample tested positive for neutralizing antibodies in the mouse neutralization assay. Additionally, relatively few subjects had multiple treatment sessions with follow up in manner to assess the longer term incidence of development of antibodies. While some further clarification of the available data will be important to provide in current labeling, the most important questions can only be answered by data not presently available.

Comment:

This question requires further inquiry into the existing data by Elan. Planning for post-marketing studies will also be needed.

SUMMARY

CLINICAL DEVELOPMENT PROGRAM

Overview

Elan Pharmaceuticals has developed a botulinum toxin Type B for use in the treatment of cervical dystonia. The proposed commercial product is manufactured at the NPF facility, while the product used in all clinical studies submitted to date was manufactured in the facility.

The clinical development program consists of 11 studies, 5 of which are phase 1 studies. There were two phase 2 studies, two phase 3 confirmatory studies, and two open label studies for expansion of the safety experience dataset. The two open label studies remained ongoing at the time of the data cutoff for this BLA submission, April 1998.

Phase 2 Studies

The two phase 2 studies, AN072-008 and AN072-009 were dose ranging studies, overlapping in dose only by a single dose level commonly examined in both studies (2400 U in Study 008, 2500 U in Study 009).

Study 008 was not reviewed in depth. Study 008 was a placebo controlled, dose ranging study of 3 doses ranging from 400 U to 2400 in a single treatment session. There are no known problematic study quality issues regarding this study preventing the results from this study being used for hypothesis generating purposes.

Study 009 was a placebo controlled, dose ranging study of 3 dose levels of 2500 U, 5000 U and 10000 U. This study appears to have been conducted in a manner consistent with GCP. The design of this study allowed for widely varying duration of subject participation due to termination of subject participation when each subject was first deemed without clinical response or retention of previously apparent response. This prohibits most study comparisons after the Week 4 evaluation.

A total of 122 subjects enrolled into Study 009 across 12 sites. Demographics and baseline treatment characteristics were similar between groups.

Phase 3 Controlled Studies

Study 301 and Study 302 were the two phase 3 confirmatory studies. They were conducted at the same time, in an overlapping set of study centers, with nearly the same design. The major two differences in design were that Study 301 compared 2 doses of BotTx-B with placebo, while Study 302 compared only 1 dose, and that Study 302 enrolled only patients who were no longer responsive to botulinum toxin Type A, while Study 301 only subjects who remained responsive to type A toxin.

In these two studies subjects received a single treatment session with the randomized assigned treatment at the start of the study and were followed every 4 weeks thereafter. Study treatment was planned for injection into 2 to 4 affected muscles, per the investigator's judgement. Study efficacy evaluations included the Toronto Western Spasmodic Torticollis Rating Scale

(TWSTRS), Patient and Investigator Global VAS Assessment of Change, and a Patient Pain VAS Assessment. The primary endpoint for both studies was the pairwise comparison of the 10000 U dose vs placebo on the TWSTRS score at Week 4.

A total of 109 subjects enrolled into Study 301, divided into 3 treatment groups; a total of 77 subjects were enrolled into Study 302, and randomized to two treatment groups. Most subjects completed study participation as planned. There was a minor error in randomization in Study 301 that does not seriously impair the study. There were no major protocol deviations in conduct reported for either study. Study 301 was conducted at 9 sites, Study 302 at 7 sites, of which 4 sites overlapped between the two studies.

The study populations were well balanced between groups, and appear to represent subjects of moderate severity of disease.

Open Label Safety Studies

Study 352 was an open label follow-on study to assess the safety of 3 repeated doses of toxin as well as the safety of higher dose levels of toxin, up to 15000 U. There were 138 subjects of the 160 planned enrolled as of the data submission cutoff date of April 1998.

Study 351 was the primary long term open label extension study. Subjects were eligible for ongoing open label toxin treatment as needed. Doses were permitted to vary as deemed appropriate to the investigator.

EFFICACY

The phase 3 confirmatory studies both showed statistically significant treatment effects on the primary endpoint of (change from baseline in) TWSTRS score at Week 4. The change from baseline for the 10000 U group in both studies was approximately 11 points compared to 4 points or 2 points in the placebo groups. The 5000 U group in Study 301 showed a statistically significant treatment effect as well, with a mean improvement from baseline of approximately 9 points. Many subjects maintained improvements for substantial amounts of time, with median time to return to baseline estimated as between 12 to 14 weeks for all toxin treated groups.

Examination of the subscales of the TWSTRS indicates that all three subscales, Severity, Disability, and Pain, contributed to the overall effect, but the contribution of the Pain Subscale was the most prominent in both studies. Treatment effects were evident at the majority of the 12 different study sites that participated between the two studies. There was considerable variability in the size of the treatment effect between the two studies at the 4 study centers that participated in both studies.

The secondary endpoint of Patient Global Assessment of Change also supported a statistically significant treatment associated benefit in both studies. The week 4 Patient Global Assessment (100 mm VAS) was approximately 60 to 64 in the three toxin treated groups, compared to approximately 40 or 44 in the placebo groups.

The tertiary efficacy endpoints in these studies also showed statistically significant treatment associated benefits, fully supporting the primary finding of efficacy.

Although Elan conducted a responder analysis, there was no prospectively agreed upon meaningful definition of a responder, and post hoc examinations of the study results suggests that the Elan criteria for a responder are seriously flawed in the subjects classified as a responder.

Overall, these two studies indicate a statistically significant treatment effect. This treatment effect is most prominent on the pain component of cervical dystonia. The mean treatment effect in size is approximately 1/4 to 1/3 of the baseline severity. Thus, while subjects are improved on their symptoms, no subjects are entirely relieved of their disease symptoms.

Efficacy of the toxin treatment was also demonstrated in Study 009. All three dose levels were shown to be efficacious. Treatment effect size had appeared to be slightly larger in this phase 2 study than was shown in the subsequently conducted phase 3 studies. TWSTRS subscales again showed contribution by all three components, and the secondary efficacy parameters were also statistically significant for all three dose levels.

Study 008 did not achieve statistical significance on it's primary efficacy endpoint. However, secondary analyses and endpoints were suggestive that the 2400 U dose was efficacious.

SAFETY

Safety results were highly consistent between the studies. There were no deaths or serious adverse events regarded as related to study treatment. The most prominently treatment associated adverse events were dry mouth and dysphagia. Both are dose related in incidence. Injection site pain also appears to have a treatment associated incidence at higher doses of toxin, and of more modest size of increased incidence. This may also be true for dyspepsia, but this is less consistently found.

Dysphagia is the most common potentially severe or serious adverse event. No subject in any of these studies experience serious dysphagia, and the few who had severe dysphagia did not require tube feeding for nutritional support or any other major changes in activities during the transient period of the adverse event. Although not serious, dysphagia is important because it has the potential, when severe, to cause marked disruption in a patient's functioning, and the potential to become a serious AE.

Dysphagia and dry mouth were also important because these events, even when only moderate in intensity, were responsible for subjects discontinuing study participation and declining further injection sessions with BotTx-B

OTHER ISSUES

Antibody formation is a potential concern for both risk of adverse events and for loss of efficacy. A two stage antibody testing procedure is employed by Elan. While a modest number of subjects developed low to moderate levels of ELISA detected antibodies, only 1 of these subjects developed neutralizing antibodies on the mouse neutralization assay. However, antibody testing

results have been reported in too unclear a manner to draw any firm conclusions on the extent of the testing. Additionally, at best the antibody testing to date is preliminary, not having had the opportunity to follow extensive numbers of subjects for extended amounts of time yet.

Exploratory analyses of the datasets indicated that there were not major disparities in efficacy between subjects based on baseline dystonia severity, sex, or age. Subjects in all subdivisions by any one of these factors all had clearly treatment associated benefit. There were too few subjects in race related subset to perform any analysis.

Comparison of the amount of efficacy between dose levels suggests that a dose of 2500 U provides significant benefit to patients, and that there is not any apparent major difference in efficacy between 5000 U and 10000 U, or between 10000 U and 15000 U. However, the datasets for these comparisons are limited, and small differences in efficacy are below the limits of what can be established by these studies.

Safety, however, had a clearly dose related incidence between 5000 U and 10000 U for the two most important AES, dysphagia and dry mouth. Thus, the risk benefit comparison is less favorable for the 10000 U dose than the 5000 U dose.

An additional issue is the production source of the toxin. All clinical experience reported upon and reviewed in this BLA was based upon toxin produced at —— However, only NPF produced toxin is proposed for marketing. Biophysical and preclinical characterization and comparison of these toxins are outside the scope of this review, and other reviews should be consulted for this information. However, at least some safety experience would be preferred with the NPF toxin prior to approval, given the clinical uncertainties regarding factors that influence the relative balance of safety and efficacy.

RECOMMENDATION

The applicant appears to have provided good evidence of a favorable risk to benefit comparison. This product should be ultimately approved for marketing for the treatment of cervical dystonia.

However, there are multiple issues that remain inadequately evaluated within the submission to be able to formulate adequate labeling at this time. The applicant should be required to supply the appropriate additional or clarified analyses of the dataset.

At the same time, there are also important issues of safety assessment and labeling that can be informed upon by the open label studies that remained ongoing at the time of submission. Some of these issues may be greatly clarified with submission of updated information from these studies, which should be requested. These questions include the relative efficacy of higher doses of toxin, and the safety of the toxin derived from the NPF facility at proposed doses are amongst these.

Additionally, the manufacturer should be requested to plan and initiate a longer term study to assess for the incidence and significance of antibody formation as a phase 4 commitment.

APPENDIX A

SYNOPSES OF PHASE 1 STUDIES

Study AN072-001

Title: A dose ranging safety evaluation of botulinum toxin serotype-B in patients with cervical dystonia (torticollis)

Study conducted 2/93 to 4/94. Report completed 11/94.

This was a single center, open label, dose escalation study of injections of BotTx-B into 2 to 4 neck and shoulder muscles in subjects with cervical dystonia. Injections could be repeated as soon as 4 weeks after the prior injection if there were not persistent improvement of adverse events. There were 8 subjects enrolled, who remained in the study between 127 to 398 days, receiving 1 to 5 injection sessions, 4 subjects receiving 4 or 5 treatment sessions. Dose per session ranged 100 U to 1200 U. Some subjects appeared to show improvement of symptoms, but no clear dose response curve was identified. The treatments were generally well tolerated, with few adverse events, and none severe or serious.

Study AN072-002

Title: An open label dose escalation safety and preliminary efficacy study of botulinum toxin serotype B (BotB) in patients with cervical dystonia (torticollis) who have become resistant to serotype A.

Study conducted 10/93 to 5/94. Report completed 2/95.

This open label dose escalation study was conducted at 3 centers who enrolled 12 subjects with cervical dystonia who had become secondarily resistant to botulinum toxin type A. Subjects received injections in to 2 to 4 neck muscles. There was both intrasubject and intersubject dose escalation employed. Repeat doses in individual subjects could be administered 2 weeks after a prior dose if no response had been observed, for a maximum of 3 dosing sessions per subject.

Individual doses ranged from 150 to 1430 U, 8 of the 12 subjects received 2 or 3 treatment sessions. There were no deaths, serious adverse events, or withdrawal due to adverse event. There was a suggestion of efficacy in the higher dose treatment sessions compared to the lower dose sessions in the TWSTRS-Severity subscale.

Study AN072-003

Title: An open label dose escalation safety and preliminary efficacy study of botulinum toxin serotype B (BotB) in patients with cervical dystonia (torticollis).

Study conducted 12/93 to 1/96. Report completed 1/97

Three centers participated in this open label, non-randomized study of subjects with cervical dystonia who received injections of toxin into neck muscles. Doses were escalated both intrasubject and intersubject, and could be repeated as soon as 2 week after the prior injection if no efficacy nor adverse effect was apparent. Only the first 8 subjects were eligible for repeat injections, the remaining 20 subjects received only 1 injection. Individual dosing sessions ranged from 300 to 12000 units. 19 doses were 2400 U or higher. Subjects were retrospectively divided into groups based on doses received. The higher doses tended to indicate greater clinical improvement response then the lower dose sessions. There were no deaths, serious adverse events, or withdrawals from study due to AE.

Study AN072-120 and AN072-121

Study 120 Title: Electrophysiologic response of the extensor digitorum brevis to varying doses of intramuscularly injected BotB in normal subjects.

Study conducted July 1994

Study 121 Title: Electrophysiologic response of the extensor digitorum brevis to varying doses of intramuscularly injected BotB (botulinum toxin type B) and BOTOX (botulinum toxin type A) in normal subjects.

Study conducted October 1994

Both of these studies were conducted by a single investigator. In Study 120, each EDB muscle in 18 healthy volunteers were randomized to receive different doses (from 17 possible choices) of Type B toxin, with one control subject receiving 0 U in both EDB muscles. Thus, each dose was injected into 2 subjects. Doses ranged from 1.25 to 480 U of toxin. EDB M-waves were assessed for activity of the toxin. Maximal suppression of the M-wave was 75%, with a dose response curve defined.

In Study 121, 11 healthy volunteers received injections of BOTOX into one EDB, and Type B toxin into the other EDB (with one subject as a control). Doses of BOTOX ranged 1.25 U to 10 U, of Type B toxin 20U to 480 U, with pairing of the dose levels between toxins. Similar but not identical amounts of decline in M-wave, and duration of effect were seen between BOTOX and Type B toxin. In general, the amount of decline for Type B toxin was less than that for BOTOX.

These two studies were combined into a single publication in the literature, Sloop RR, et.al., 1997, Neurology 49:189-194.